

STIC-Biotech/ChemLib

72632

From: Rao, Manjunath N.
Sent: Tuesday, August 06, 2002 3:09 PM
To: STIC-Biotech/ChemLib
Subj ct: Sequence search request for 09/742,690

From: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10C 01
Phone: 306-5681

Date: 8-6-02

Please search the following as soon as possible for application with serial number **09/742,690**

SEQ ID NO: 1 against all commercial nucleic acid databases including issued patents database and pending application database and provide a print of **all results**.

SEQ ID NO: 1 against all commercial protein databases including issued patents database and pending application database and provide a print of **all results**.

If you have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.
Biotechnology Patent Examiner
Art Unit 1652, Room 10A11
Crystal Mall 1, USPTO.

OFFICE OF THE
PATENT EXAMINER
AUG 14 2002

Point of Contact
P. Sheppard
Searcher: Telephone number (703) 308-4499
Phone: _____
Location: _____
Date Picked Up: _____
Date Completed: 8/14/02
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 06:44:47 ; Search time 1772.03 Seconds

(without alignments)
4250.098 Million cell updates/sec

Title: US-09-742-690-1

Perfect score: 558

Sequence: 1 tcgagaaaacacaggtgcag.....ctccagtgctgtaataag 558

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	199.2	35.7	1071	10	BM454450
2	195.2	35.0	663	10	AGENCOURT
3	194.4	34.8	1164	10	BF974771
4	191.8	34.4	669	10	BF974771
5	189.2	33.9	687	10	BF974771
6	189.2	33.9	832	10	BF974771
7	187.6	33.6	518	9	AG402029
8	185.8	33.3	802	10	BF974771
9	185.2	33.2	571	10	BF974771
10	182.8	32.8	648	10	BF974771
11	182.4	32.7	559	9	AW401799
12	182.4	32.7	759	10	BF974771
13	182.2	32.7	493	9	AW402069
14	182	32.6	471	9	AW403220
15	182	32.6	516	9	AW401728
16	181.4	32.5	800	10	BM007901
17	180.6	32.4	874	10	BI755805

18	180.4	32.3	445	9	AW408316
19	179.6	32.2	846	10	BF755572
20	179.2	32.1	705	10	BF755572
21	178.8	32.0	750	10	BF976111
22	178.6	32.0	503	9	AW402572
23	178.2	31.9	735	10	BF745292
24	177.8	31.9	487	9	AW403862
25	177.6	31.8	695	10	BF757738
26	177.2	31.8	870	10	BF757738
27	177.2	31.8	908	10	BF757738
28	177	31.7	396	9	AW407843
29	177	31.7	912	10	BF756211
30	176.2	31.6	930	10	BF756211
31	175.8	31.5	717	10	BF754232
32	175.6	31.5	412	9	AW402895
33	175.6	31.5	595	10	BF759526
34	174.8	31.3	940	10	BF754180
35	174.6	31.3	740	10	BF757947
36	174.6	31.3	741	10	BF754024
37	174.4	31.3	511	9	AW402613
38	174.4	31.3	678	10	BF754064
39	174.4	31.3	870	10	BF759649
40	174.4	31.3	991	10	BF974524
41	174.2	31.2	536	9	AW402624
42	174.2	31.2	769	10	BM007986
43	174	31.2	491	9	AW402907
44	174	31.2	704	10	BF759119
45	173.8	31.1	483	9	AW403901

ALIGNMENTS

RESULT 1

BM454450

LOCUS

DEFINITION

5', mRNA sequence.

BM454450

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM454450 1071 bp mRNA linear EST 05-FEB-2002
AGENCOURT_6406467 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5582873

BM454450.1 GI:18503490

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1. (bases 1 to 1071)

NIH-MGC http://mgi.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaaps-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL2345 row: a column: 18

High quality sequence stop: 425.

Location/Qualifiers

1..1071

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5582873"

/clone_lib="NIH_MGC_92"

/tissue_type="embryonal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="Organ: testis; Vector: pCMV-SPORT6; Site: 1: NotI;

Site: 2: SalI; Cloned unidirectionally; oligo-dr primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

```

BASE COUNT      195 a      304 c      303 g      269 t
ORIGIN

Query Match      35.7%; Score 199.2; DB 10; Length 1071;
Best Local Similarity 67.9%; Pred. No. 2.2e-48;
Matches 326; Conservative 0; Mismatches 143; Indels 11; Gaps 3;

QY 13 aggtgcagctgcagagctcaggggagagattggtgcaggctgggggctctctgagactct 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 221 AGGTGCAGCTGGCGAGCTCTGGGGAGAGTGTGTACGGCTGGGGGCTCCCTGAGACTCT 280

QY 73 cctgtgcagctcgggcgcgcgcacccagtggtcatgtcactatggtatgggtgggtcc 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 281 CCTTGCAGCCTCTGGATTCCAC-----TTTGATGATTATGGCATGAGCTGGGTCC 331

QY 133 gccaggttcagggaagagcgtagtctgtcgcagctattagtgagtggttaagaga 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 332 GCCAAGCTCCAGGAAGGGCTGAGTGGGTCTGTGTATTAATGGAATGGTGGCTAGCA 391

QY 193 catggtataaagactcgtgaaggccgattccaccatctccagagataaacgcaagacta 252
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 392 CAGGTTATGCAGACTCTGTGAAGGGCGGATTCCACCATCTCCAGAGACAAGCCCAAGAACT 451

QY 253 cgggttattcgaataaagacagctgaacctgaagatacagccgtttattattgtgc-c 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 452 CCCTGTATCTGCAANTGAACAGCCCTGAGAGTCGAGGACACGGCCCTTGTACACTGTGCGA 511

QY 312 gctgcacggtccgcgtggatgatattccctgcgcgttggtgttgactactggtggccag 371
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 512 GAGATTACAGTACTACTAATGGTGTATGCTATACGGACTGGTTTCGACCCCTCGGGCCAG 571

QY 372 gggaccaggtcacctctctcctcagg-atctcatcaccatcaccatcacggatccacctc 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 572 GGAACCTGGTTCACCGTCTCTCAGGGCGAGTCTTCACACACCCCTCTCTGAGTCCACTTA 631

QY 431 cattgaaggtctgtaccagctcactacggtcaggtcagtggtggtattggttactccggtcc 490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 632 GGGAGATCCCTTGCACAGGGTCTCAGGGTTACAGTCTCCTGGAGGCCATTTCGAGGTCC 691

RESULT 2
BG686759      663 bp      mRNA      linear      EST 01-MAY-2001
LOCUS      602650729F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4763215 5',
DEFINITION      mRNA sequence.
ACCESSION      BG686759
VERSION      BG686759.1 GI:13918156
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 663)
AUTHORS      NTH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LICM1618 row: i column: 08
            High quality sequence stop: 659.
            Location/Qualifiers
            1. 663
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4763215"

FEATURES
source

```

```

/clone_lib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: B-cells; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GCCACGAG(G). Size selected by Ling
for average insert size 1.8kb. Library constructed by Ling
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

BASE COUNT      138 a      185 c      186 g      154 t
ORIGIN

Query Match      35.0%; Score 195.2; DB 10; Length 663;
Best Local Similarity 71.1%; Pred. No. 2.7e-47;
Matches 276; Conservative 0; Mismatches 103; Indels 9; Gaps 1;

QY 10 gacaggtgcagctgcagagtcaggggagagattggtgcagcgtgggggctctctgagac 69
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 126 GTCAGGTGCAGCTGCTGGAGTCTGGGGGAGGCGTGGTCCAGCCTGGGAGGTCCCTGAGAC 185

QY 70 tctctatgcagctcggagcgcgcaccagtgctcatgtcactatggtatggctggt 129
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 186 TCTCCTGTGACGCTCTGGATTCCACTTCAGTAG-----CTATGCTATGCACTGGG 236

QY 130 tccgccaggttccagggaagcgtgagttgtcgcagctattagtggtgagtggtgaaag 189
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 237 TCCGCCAGGCTCCAGGCAAGGGGCTGGAGTGGGTGGCATTTATATCATATGATGGAAGCA 296

QY 190 agacatggtataaagactccgtgaaggccgattcaccatctccagagataaacgcaaga 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 297 ATAATAACTAGCGACACTCCGTGAAGGCCGATTCCACCATCTCCAGACACAATTCACAGA 356

QY 250 ctacggttattctcgaatgaacagcctgaacctgaagatacagccgtttattattgtg 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 357 ACAGCGTATCTGCCAATGAACAGCCCTGAGAGCTGAGACACGCGCTGTATTAFTACTGTG 416

QY 310 ccgctcagccggtccgcgtggatgatatttccctcccggttgggtttgactactggggcc 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 417 CGAGAGATCTTTTTCGTATAGCAGTGGCTGGCCATCGGTACCCCTCGCTACTGGGGCC 476

QY 370 aggggacccaggtcacccgtctcctcagg 397
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 477 AGGGAACCTGTGCACCGTCTCCTCAGG 504

RESULT 3
BF974771
LOCUS      602245420F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:436541 5',
DEFINITION      mRNA sequence.
ACCESSION      BF974771
VERSION      BF974771.1 GI:12341986
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1164)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:

```

http://image.llnl.gov
 Plate: LNCM1209 row: 9 column: 06
 High quality sequence stop: 696.
 Location/Qualifiers
 1. .1164

FEATURES

source

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4336541"
 /clone_lib="NIH_MGC_48"
 /tissue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;
 Site:2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 BASE COUNT 284 a 313 c 330 g 236 t 1 others
 ORIGIN

Query Match 34.8%; Score 194.4; DB 10; Length 1164;
 Best Local Similarity 72.8%; Pred. No. 6.3e-47;
 Matches 287; Conservative 0; Mismatches 86; Indels 21; Gaps 2;

Qy 7 aaagacaggtcacgtcagagtcagggcagagattggtcagcgtggggtctctga 66
 Db 116 AATGTCAGGTGCAGCTGTGTGAGCTGGGGAGGCTTGTCAAGCTCGAGGTCCTCA 175
 Qy 67 gactctctgtcagctcgggagcgcaccagtggtcactatggtatgggct 126
 Db 176 GACTCTCTGTGCAGCCTCTGGATTCACTTCAGTG-----ACTACTACATGAGCT 226
 Qy 127 ggttcgcagagttccagggagagcgtgagttgtcgcagctattaggtagtgta 186
 Db 227 GGATCCGCAGGCTCCAGGAGAGGGGCTGGAGTGGGTTTCATACATTAGTAGTAGTA 286
 Qy 187 aagagacaggtataaaagactccgtgaaggccgattccacatctccagagataacgca 246
 Db 287 GTTACAAAGCTACCGAGACTCTGTGAAGGCGCGATTACCATCTCCAGAGACAAGCCA 346
 Qy 247 agactacggtttatctgcaaatgaacagcctgaacacctgaagatacgcgcgtttatt 306
 Db 347 AGAACTACTGTATCTGCAAAATGAACAGCCTGAGAGCCGAGGACACGCGCGTATTACT 406
 Qy 307 gtgcgcgtcgcagcgtccgcgtgagatgatatctccctgcgcggttggttgactactgg 366
 Db 407 GTGC-----GAGAGATCAGAGGACAACACGAGCTGTAGATTTTGACTACTGGG 454
 Qy 367 gccagggaccaggtcacctcctcctcaggatc 400
 Db 455 GCCAGGGAACCTGTGTACCGTCTCCTCAGCCTC 488

RESULT 4
 LOCUS BG686716 669 bp mRNA linear EST 01-MAY-2001
 DEFINITION 602650678F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4763274 5',
 mRNA sequence.

ACCESSION BG686716
 VERSION BG686716.1 GI:13918113
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 669)
 REFERENCE NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: rgaabs-r@mail.nih.gov
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LNCM1618 row: k column: 19
 High quality sequence stop: 665.

FEATURES

Location/Qualifiers

source

1. .669
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4763274"
 /clone_lib="NIH_MGC_48"
 /tissue_type="primary B-cells from tonsils (cell line)"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;
 Site:2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Size-selected >500bp
 for average insert size 1.8kb. Library constructed by Ling
 Hong in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 BASE COUNT 149 a 184 c 186 g 150 t
 ORIGIN

Query Match 34.4%; Score 191.8; DB 10; Length 669;
 Best Local Similarity 72.1%; Pred. No. 2.8e-46;
 Matches 282; Conservative 0; Mismatches 97; Indels 12; Gaps 2;

Qy 10 gacagtcagctcagagtcagggagagattggtcagcgtggggtctctcagac 69
 Db 125 GTCAGTGCAGCTGTGTGAGCTTGGGGAGGCTTGTCAAGCCTGGAGGTCCTCGAGAC 184
 Qy 70 tctctgtcagcctcgggagcgcaccagtggttcactatggttgagtggtgaaag 129
 Db 185 TCTCTGTGCAGCCTCTGGATTCACTTCAGTG-----ACTACTACATGAGTGGA 235
 Qy 130 tccgcagagttccagggagagcgtgagttgtcgcagctattaggtagtggtgaaag 189
 Db 236 TCCGCCAGGCTCCAGGAGAGGGGCTGGAGTGGGTTTCATACATGACTAGTGGTAGTA 295
 Qy 190 agacatggtataaaagactccgtgaaggccgattccaccatccagagataaacgccaaga 249
 Db 296 TCATAAAGTACCGAGACTCTGTGAAGGGCGGATTCACCATCTCCAGGGACACGCCAAGA 355
 Qy 250 ctacggtttatctgcaaatgaacagcctgaacacctgaagatacgcgctttatttggtg 309
 Db 356 AGTCACTGTATCTGCAAAATGAACAGCCTGAGAGCCGAGGACACGCGCGTATTACTGTG 415
 Qy 310 ccgctcgcagcgttcgcgctggtgagtg----atattccctgcgcggttggttgactactgg 366
 Db 416 CGAGAGGGCAGCCCTTCAGGAACAGCTGAATATAGTGGCTACGAGATACTTTGACTACTGGG 475
 Qy 367 gccagggaccaggtcacctcctcctcagg 397
 Db 476 GCCAGGGAACCTGTGTACCGTCTCCTCTCAGG 506

RESULT 5

BG696390

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

BG696390 687 bp mRNA linear EST 07-MAY-2001
 602658778F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4802059 5',
 mRNA sequence.
 BG696390
 BG696390.1 GI:13961484
 EST.

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 687)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
 Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10695 row: k column: 20
 High quality sequence stop: 687.
 Location/Qualifiers
 1..687
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4802059"
 /clone_lib="NCI CGAP_Skn3"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI CGAP Library."
 143 a 188 c 207 g 149 t

BASE COUNT 143 a 188 c 207 g 149 t
 ORIGIN

Query Match 33.9%; Score 189.2; DB 10; Length 687;
 Best Local Similarity 70.0%; Pred. No. 1.7e-45;
 Matches 273; Conservative 0; Mismatches 108; Indels 9; Gaps 1;

QY 13 aggtgcagctgcagagtcaggggagagattggtgcagctggggctctctgagactct 72
 DB 103 AGGTGCAGTGTGGAGTCTGGGGAGGCTTGGTACAGCGGGGGTCCCTGAGACTCT 162
 QY 73 cctgtgcagctcggagcgcagccagctggtcagctattaggtgagtggtgctgtcc 132
 DB 163 CCTGTGCAGCTCTGAGTTCGGCTTTAGTAGT-----TATGCCATGAGCTGGTCC 213
 QY 133 gccaggttcagggaagagcgtgattgtgcagctattaggtgagtggtgtaaagaga 192
 DB 214 GCCAGGCTCCAGGAGGGCTGGAGTGGTCTCGGCCATTTAGTGTCTGTGGCACCA 273
 QY 193 catggtataaagactccgtgaaggccgattccaccatctccagagataaacgcaagacta 252
 DB 274 CATACTATATAGACTCCGTGAAGGCGCGTTTCATTATCTCCAGAGACAATTCAGAAC 333
 QY 253 cgggtttatctgaaatgaacagcctgaacacctgaagatacggccgtttattattgtgcg 312
 DB 334 CGTTGTATCTGCAATATGAACAGCCTGAGAGCGGCAACACAGCCCATATATTACTGTGCGA 393
 QY 313 ctgcagccgtccgcgtggtgatatttccctgcggtggtttgaactactgggcccagg 372
 DB 394 AATTAGCCGCTGGACAGTACAGCTGTGCTAGTATTATACGGCATGGAGGCTGGGCCAAG 453
 QY 373 ggaccaggtccagctctcctcaggatctc 402
 DB 454 GGACCAGGTCACCGTCTCTCAGCATCCC 483

RESULT 6
 BG697050 832 bp mRNA linear EST 07-MAY-2001
 LOCUS 602660253F1 NCI CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4803540 5',
 DEFINITION mRNA sequence.
 ACCESSION BG697050

VERSION BG697050.1 GI:13962837
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 832)
 AUTHORS NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
 Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10699 row: i column: 13
 High quality sequence stop: 830.
 Location/Qualifiers
 1..832
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4803540"
 /clone_lib="NCI CGAP_Skn3"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI CGAP Library."
 174 a 250 c 232 g 176 t

BASE COUNT 174 a 250 c 232 g 176 t
 ORIGIN

Query Match 33.9%; Score 189.2; DB 10; Length 832;
 Best Local Similarity 70.0%; Pred. No. 1.9e-45;
 Matches 273; Conservative 0; Mismatches 108; Indels 9; Gaps 1;

QY 13 aggtgcagctgcagagtcaggggagagattggtgcagctggggctctctgagactct 72
 DB 102 AGGTGCAGTGTGGAGTCTGGGGAGGCTTGGTACAGCGGGGGTCCCTGAGACTCT 161
 QY 73 cctgtgcagctcggagcgcagccagctggtcagctattaggtgagtggtgtaaagaga 192
 DB 162 CCTGTGCAGCTCTGAGTTCGGCTTTAGTAGT-----TATGCCATGAGCTGGTCC 212
 QY 133 gccaggttcagggaagagcgtgattgtgcagctattaggtgagtggtgtaaagaga 192
 DB 213 GCCAGGCTCCAGGAGGGCTGGAGTGGTCTCGGCCATTTAGTGTCTGTGGCACCA 272
 QY 193 catggtataaagactccgtgaaggccgattccaccatctccagagataaacgcaagacta 252
 DB 273 CATACTATATAGACTCCGTGAAGGCGCGTTTCATTATCTCCAGAGACAATTCAGAAC 332
 QY 253 cgggtttatctgaaatgaacagcctgaacacctgaagatacggccgtttattattgtgcg 312
 DB 333 CGTTGTATCTGCAATATGAACAGCCTGAGAGCGGCAACACAGCCCATATATTACTGTGCGA 392
 QY 313 ctgcagccgtccgcgtggtgatatttccctgcggtggtttgaactactgggcccagg 372
 DB 393 AATTAGCCGCTGGACAGTACAGCTGTGCTAGTATTATACGGCATGGAGGCTGGGCCAAG 452
 QY 373 ggaccaggtccagctctcctcaggatctc 402
 DB 453 GGACCAGGTCACCGTCTCTCAGCATCCC 482

RESULT 7
 BG697050 832 bp mRNA linear EST 16-FEB-2000
 LOCUS 602660253F1 NCI CGAP_Skn3 Homo sapiens cDNA clone
 DEFINITION UI-HF-BK0-aap-a-04-0-UI.r1 NIH_MGC_36 Homo sapiens cDNA clone

IMAGE:3054247 5', mRNA sequence.

AW402029
VERSION
KEYWORDS
SOURCE
ORGANISM

Human sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 518)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: M.B. Soares Lab
CDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://bio.llnl.gov/bbrp/image/image.html>
Seq primer: M13 Forward.

FEATURES
source

Location/Qualifiers
1..518
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3054247"
/clone_lib="NIH_MGC_36"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/cell_line="MGC85"
/lab_host="DH10B (lrr)"
/note="Vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(0.5-1.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonafide, Ph.D. and M. Bento Soares, Ph.D."
99 a 149 c 148 g 121 t 1 others

Query Match 33.6%; Score 187.6; DB 9; Length 518;
Best Local Similarity 73.7%; Pred. No. 4.5e-45;
Matches 286; Conservative 0; Mismatches 84; Indels 18; Gaps 3;

Qy 13 agtgcagctgcaggagtcaggagagattggtgcagctggggctctgcagactc 72
Db 54 ACgtgcagctgcaggagtcaggagagattggtgcagctggggctctgcagactc 113

Qy 73 cctgtcagctgcaggagtcaggagagattggtgcagctggggctctgcagactc 132
Db 114 Cctgtgcagctgcaggagtcaggagagattggtgcagctggggctctgcagactc 164

Qy 133 gccaggttccaggagagagtcaggagagattggtgcagctggggctctgcagactc 192
Db 165 GCCAGGCTCCAGGAGAGGCTGGAGTGGGTCTCAATTATTAGTGTACTAGCA 224

Qy 193 catggtataaagactccgtgaaggccgattaccatctccagagataacgccaagacta 252
Db 225 CATACATCCAGCAGACTCCGTGAAGGCGGTTCATCATCTCCAGAGACAATCCAA 284

Qy 253 cgggttatctgcaaatgaacagcctgaacacctgaagatacgcggtttattattgtgcg 312
Db 285 CGCTGTCTTGCAGAAATGAACAGCTGAGAGTCCGAGGACAGCCGCTATATTACTGTGC 342

Qy 313 ctogaccggtccgctggatgatacttccctgcgcggttgggttggactactggggccagg 372
Db 343 --GNAAGGCTCACAGTGGCTGA-----CCOCTCTGCTCTTGACTACTGGGCCACG 395

Qy 373 ggaccaggttcacgtctctcaggatc 400

Db 396 GTACCCCTGGTCAACGCTCTCTCTCAGCCTC 423

RESULT 8
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Human sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 802)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLC1700 row: a column: 16
High quality sequence stop: 795.
Location/Qualifiers
1..802
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4853655"
/clone_lib="NIH_MGC_48"
/tissue_type="primary B-cells from tonsils (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; CDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCAGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by Lin
Hong in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: This is a NIH_MGC library."
180 a 221 c 234 g 167 t

Query Match 33.3%; Score 185.8; DB 10; Length 802;
Best Local Similarity 71.2%; Pred. No. 1.9e-44;
Matches 277; Conservative 0; Mismatches 102; Indels 10; Gaps 2

Qy 10 gacaggtgcagctgcaggagtcaggagagattggtgcagctggggctctctctgagac 69
Db 126 GTCAGGTGCAGCTGGTGGAGTCTGGGGAGGCGGTGCTCCAGCTGGGGGTCTCCCTGAGAC 185

Qy 70 tctcctgtcagctgcaggagtcaggagagattggtgcagctggggctctctctgagctggt 129
Db 186 TCTCCTGTGAGCGCTCTGGGATTCACCTTCCTAGTAG-----CTATGGCATGCATCTGGG 236

Qy 130 tccgccaggttccaggagagtcaggagagattggtgcagctattagtgagtggtgta-aa 188
Db 237 TCCGCCAGGCTCCAGGCAAGGGCTGGAGTGGGTGGCATTTATACGTTATGATGGAACGCT 296

Qy 189 gagacatggtataaagactccgtgaaggccgattaccatctccagagataacgccaag 248
Db 297 AATAAATACATGCAGACTCCGTGAAGGGCCGATTCCACCTCTCCAGAGACAATCCCAAG 356

Qy 249 actacggtttatctgcaaatgaacagcctgaacacctgaagatacgcgctttattattgt 308
Db 357 AACACGCTGTATCTGCAAAATGAACAGCTGAGAGCTGAGACACGCTGTGTATTACTGT 416

```

Qy 309 gccgcgcagcgcgtccgcggtgatattttccctgcgcggttggtggttgaactactcagcgcgcgc 368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 417 CGAAGAAGAGAGTGGGAGCTCCCTGCGGTACTACTACTATGTGTATGGACGTCTGGGCG 476

Qy 369 caggggaccaggtccagctctcctcagg 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 477 CAAGGACACGGTCACCGCTCCCTCAGG 505

RESULT 9
LOCUS BG759713 571 bp mRNA linear EST 15-MAY-2001
DEFINITION 602711110F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4851387 5',
    mRNA sequence.
ACCESSION BG759713
VERSION BG759713.1 GI:14070366
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
    Email: cgapbs-r@mail.nih.gov
    Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
    cDNA Library Preparation: Ling Hong/Rubin Laboratory
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LLNL at:
    http://image.llnl.gov
    Plate: L16M1694 row: c column: 04
    High quality sequence stop: 569.
    Location/Qualifiers
        1. 571
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4851387"
            /clone_lib="NIH_MGC_48"
            /tissue_type="primary B-cells from tonsils (cell line)"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;
                Site:2: EcoRI; cDNA made by oligo-dt priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                for average insert size 1.8kb. Library constructed by Ling
                Hong in the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH_MGC Library."
                BASE COUNT 119 a 165 c 160 g 127 t
                ORIGIN
                Query Match 33.2%; Score 185.2; DB 10; Length 571;
                Best Local Similarity 69.0%; Pred. No. 2.4e-44;
                Matches 272; Conservative 0; Mismatches 113; Indels 9; Gaps 1;

Qy 7 aaagacaggtgcagctgcagagtcaggggagagattggtgcagcgtggggctctctga 66
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 AATGTGAGCTGAGCTGTGGAGTCTGGGGAGGCTTGCTACAGCCTGGGGGTCCTCTGA 138

Qy 67 gactctctgtgacgtctcgagcgcgcaccagtggttcacatggtatggtgact 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 GACTCTCTGTGGCCCTCTGATTCACC-----TTTAGCAACATATGCCATGAGCT 189

Qy 127 ggttcgcaggttccagggagagcgtgagttgttcgcagctattagtgaggatggtga 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 190 GGGTCCGCCAGACTCCACTGACGGGGCTGGAGTGGGTCTCAGCCATTACTGACAAATGTTG 249

```

```

Qy 187 aagagacatggtatataaagactcgtgaaggccgagattccaccatctccagagataacgcga 246
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 TCAACACTTATCATCCAGACTCCGTCAAGGCGCGCTTACCATCACAGAGACAATCC 309

Qy 247 agactacggtttatctgcaaatgaacacgcctgaacacctgaagatcagccggtttattatt 306
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 310 AGAACACCCCTGTATCTGAACATGAAGAGCCTGAGAGTCGAGGACACGCGCGGTATATTACT 369

Qy 307 gtgcgcgtcgcagcgcgtccgcggtggatgatatttccctgcgcggttggttgggttgaactagg 366
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 GTGCGAAATCCCAATACGAACCTTCTACTAGTGTGAGGTGTTCACCGCTTGTGACGACTGG 429

Qy 367 gccaggggaccagcaggtccacgctctcctcaggatc 400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 GCCAGGGAACCCCTGGTCCACCGCTCTCTCAGCCTC 463

RESULT 10
LOCUS BG340670 648 bp mRNA linear EST 27-FEB-2001
DEFINITION 602462250F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4574941 5',
    mRNA sequence.
ACCESSION BG340670
VERSION BG340670.1 GI:13147108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
    Email: cgapbs-r@mail.nih.gov
    Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
    cDNA Library Preparation: Ling Hong/Rubin Laboratory
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LLNL at:
    http://image.llnl.gov
    Plate: L16M1285 row: d column: 14
    High quality sequence stop: 648.
    Location/Qualifiers
        1. 648
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4574941"
            /clone_lib="NIH_MGC_48"
            /tissue_type="primary B-cells from tonsils (cell line)"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: B-cells; Vector: pOTB7; Site:1: XhoI;
                Site:2: EcoRI; cDNA made by oligo-dt priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                for average insert size 1.8kb. Library constructed by Ling
                Hong in the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH_MGC Library."
                BASE COUNT 148 a 178 c 188 g 132 t
                ORIGIN
                Query Match 32.8%; Score 182.8; DB 10; Length 648;
                Best Local Similarity 77.1%; Pred. No. 1.4e-43;
                Matches 239; Conservative 0; Mismatches 62; Indels 9; Gaps 1;

Qy 13 aggtgcagctgcagagtcaggggagagattggtgcagcgtggggctctctgagactct 72
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 AGGTCCACCTGTTGGAGTCTGGGGAGGCTTGCTACAGCCTGGGGGTCCTCTGAGACTCT 142

```

```

QY 73 cctgtcagcctcgggacgcgcacacagtggtcatggtcactatggtatggtggtggtcc 132
Db 143 CCTGTGAGCCTGCTGATTCACCTTTAGCAG-----CTATGCCATCAGCTGGGTCC 193
QY 133 gccaggttcaggagagcgclgagtttgcgcagctattagtggtggtggtggtggtc 192
Db 194 GCCAGGCTCAGGAGAGGCGCTGAGTGGTCTCAGCTATTAGTGGTGTAGCTGGGTAGCA 253
QY 193 catggtataaagactcctgtaagggcgattccaccatctccagagataaaccgcaagacta 252
Db 254 CATACTACGACGACTCCGTGAAGGCGGTTCCACATCTCCAGAGACAATTCGAAGAACA 313
QY 253 cggttatctgcaaatgaacagcctgaaactgaagatacggcgtttattattgtgcgg 312
Db 314 CGTGTATCTGCAATGAACAGCCTGAGACCGGAGACACGGCGGTATATTACTGTGCGA 373
QY 313 ctgcaccggt 322
Db 374 CCCACCGGT 383

RESULT 11
AW401799
LOCUS
DEFINITION
UT-HE-BK0-aag-g-06-0-UI.r1 NIH_MGC_36 Homo sapiens cDNA clone
IMAGE:3053770 5', mRNA sequence.
ACCESSION
AW401799
VERSION
AW401799.1 GI:6920485
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 559)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward.
FEATURES
Location/Qualifiers
1..559
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3053770"
/clone_lib="NIH_MGC_36"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (LTI)"
/notes="Vector: p7T3-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(0.5-1.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT
116 a 150 c 158 g 135 t
ORIGIN

Query Match 32.7%; Score 182.4; DB 9; Length 559;
Best Local Similarity 69.1%; Pred. No. 1.7e-43;
Matches 268; Conservative 0; Mismatches 111; Indels 9; Gaps 1;

QY 13 aggtgcagctgcagagtcaggaggaggttggtgcaggtgggggtctctgagactct 72

```

```

Db 128 AGTGCAGCTGGTGGCGTCTGGGGGACGCCTGGTACGGCTGGGGGTCCCTCGAGACTCT 187
QY 73 cctgtgcagcctcgggacgcgcacacagtggtcatggtcactatggtggtggtggtcc 132
Db 188 CCTGTGTAGTCTCTGGATTCACT-----TTTGATGATTATGTCATGAGCTGGGTCC 238
QY 133 gccaggttcaggagagcgclgagtttgcgcagctattagtggtggtggtggtggtc 192
Db 239 GCCAAGTTCAGGAGAGGCGCTGGAGTGGGTCTCTGCCATTAAATTGGAATGGTGGCACTA 298
QY 193 catggtataaagactcctgtaagggcgattccaccatctccagagataaaccgcaagacta 252
Db 299 CAGCTTATGACGACTCTGTGAAGGCGGATTCCACATCTCCAGAGACAACGCCAAGAACT 358
QY 253 cggttatctgcaaatgaacagcctgaaactgaagatacggcgtttattattgtgcgg 312
Db 359 CCTATTCTGCAATGAACAGCTCTGAAGGCGGAGACACGGCTTCTATTACTGTGCGA 418
QY 313 ctgcaccggtcgcgtggtgatatttccctgcgcgttgggtttgactactggggccagg 372
Db 419 GAGTGGCACCAGGCTATGACAGTGCCAAACTTTACTACTTTGACTACTGGGCCAGG 478
QY 373 ggaccaggttcacgcgtctcctcaggatc 400
Db 479 GAACCTGTGTCACCGTCTCCTCAGCCTC 506

RESULT 12
BG745331
LOCUS
DEFINITION
60272365F1 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:4850125 5',
mRNA sequence.
ACCESSION
BG745331
VERSION
BG745331.1 GI:14055984
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 759)
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1690 row: n column: 14
High quality sequence stop: 759.
FEATURES
Location/Qualifiers
1..759
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4850125"
/clone_lib="NIH_MGC_113"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT
158 a 215 c 215 g 171 t
ORIGIN

```


	Query Match	32.7%; Score 182.4; DB 10;	Length 759;
	Best Local Similarity	69.1%; Pred. No. 1.9e-43;	
	Matches 268; Conservative	0; Mismatches 111; Indels	9; Gaps 1;
QY	13 aggtgcagctcaggagtccagggggagagattggtcgacgctgggggctctctcgactctt	72	
DB	128 AAGTGACAGCTGGTAGTTCGCGGAGGCITGTGTACAGCCGTCGCAAGGTCCCTGAGACTCT	187	
QY	73 cctgtgcagtcctgggaacgcgcaccagtgatcatggtcaactatggttatggcggtttccc	132	
DB	188 CCTGTGACGCTCTGGATTTCACC-----TTTGTATGATTATGCCATGCATCGGTGCC	238	
QY	133 gccaggttccaggaaagagcgtgagtgttcgccagctattagtagtgagtgtaagaaga	192	
DB	239 GGCAAGCTCCAGGAAGGGCCTGGAGTGGGTCTCAGGTATTAGTTGGAATAAGTGGTAGCA	298	
QY	193 catggtataaagactcogtgaaggccgatccaccttcacagatatccagagataaacgccagacta	252	
DB	299 CAGCGTATTGGGACTCTGTGTAAGGGCCGATTTCACCATCTCCAGAGAACGCCAAGAACA	358	
QY	253 cggtttatctgcaaatgaacagcctgaacctgaagatacagccggttattattgtgcgcg	312	
DB	359 CCCTGTATCTGCAANTGAACAGCTGTAGAGCTGAGGACAGCCGCTGTATTACTGTCAA	418	
QY	313 ctgcagccgtccgcgtgtagatattccccctgcgcgttgtggtttgactactggggccagg	372	
DB	419 AAGATTGCGCTCTACGATTTTTGGAAACAGAGGGGTGCTTTTGATATCTCGGGCCAAG	478	
QY	373 ggaccaggtcacogtctoctcaggatc	400	
DB	479 GGCAATGGTTCACCGCTCTTTCAGCCCTC	506	
 RESULT 13			
AW402069			
LOCUS UI-HF-BKO-aap-e-03-0-U1.r1 NIH_MGC_36 Homo sapiens linear EST 16-FEB-2000			
DEFINITION IMAGE:3054437 5', mRNA sequence.			
ACCESSION AW402069			
VERSION AW402069.1 GI:6920755			
KEYWORDS EST,			
SOURCE human,			
ORGANISM Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/.			
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)			
JOURNAL Unpublished (1999)			
COMMENT Contact: Robert Strausberg, Ph.D. Email: gcapsb-re@mail.nih.gov Eco RI site shown at the beginning of the sequence. Tissue procurement: Louis M. Staudt, M.D., Ph.D. CDNA Library Preparation: M.B. Soares Lab DNA Sequencing by: M.B. Soares Lab Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www.bio.lnl.gov/bbrp/image/image.html Seq primer: M13 Forward.			
FEATURES			
source			
Location/Qualifiers			
1..493			
/organism="Homo sapiens"			
/db_xref="taxon:9606"			
/clone="IMAGE:3054437"			
/clone_lib="NIH_MGC_36"			
/tissue_type="lymph"			
/cell_type="germinal center B cells"			
/cell_line="MGCS5"			
/lab_host="DH10B (LTI)"			
/note=Vector: pTFr3-Pac; Site_1: NotI; Site_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA			

```

/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="MGC85"
/lab_host="DH10B (LTI)"
/notes="Vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(0.5-1.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT      98 a   127 c   143 g   103 t
ORIGIN

Query Match      32.6%; Score 182; DB 9; Length 471;
Best Local Similarity 71.4%; Pred. No. 2e-43;
Matches 277; Conservative 0; Mismatches 90; Indels 21; Gaps 2;

QY 13 aggtgcagctgcagagtcagagggagagattggtgcagggctggggctctctgcagactct 72
DB 77 AGGTGCAGCTGGTGGAGTCTGGGGGAGGCGCTGGTCAAGCCTGGGGGTCCCTGAGACTCT 136
QY 73 cctgtgcagcctggagcggccaccagtggtcactatggtcactatggtatggctgggttcc 132
DB 137 CCTGTGAGCGCTCTGGATTCACTTTCAGTAG-----CCATAGCATGAACCTGGGTCC 187
QY 133 gccaggttccaggaagagagcgtgaattgtgcagctattagtgagtgtaagaga 192
DB 188 GCCAGGCTCCAGGAAGGGGCTGGAGTGGGTCTCATCATTTAGTAGTAGTGGTTTCA 247
QY 193 catggtataagactccgtgaagggcgattccaccatccagagataacccaagacta 252
DB 248 TATACTACGAGACTCAGTGAAGGGCGGATTCACTTCAGAGACACACCCAGACT 307
QY 253 cggtttctcgaatgaacagcctgaacccgaagatacagcggctttattattgtgcg 312
DB 308 CACTGTATCTCAAAATGAACAGCCTGAGAGCGGAGGACACGCGTGTGTATTACTGTGC-- 365
QY 313 ctgcagcgtccgctgagatatttccctgcggctgggttggtaactactggggccagg 372
DB 366 -----GAGAGCCCATAGTAGAGCTGGGCCAGCAGCAGCTACTACTGGGGCCAGG 415
QY 373 ggaccaggtcacgctctctccaggatc 400
DB 416 GAACCCCTGGTCACCGTCTCCTCAGCCCTC 443

RESULT 15
AW401728      516 bp      mRNA      linear      EST 16-FEB-2000
LOCUS      UI-HF-BK0-aaf-f-12-0-UI.r1 NIH_MGC_36 Homo sapiens cDNA clone
DEFINITION      IMAGE:3053711 5', mRNA sequence.
ACCESSION      AW401728
VERSION      AW401728.1 GI:6920414
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 516)
AUTHORS      NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Eco RI site shown at the beginning of the sequence.
cDNA Library Preparation: M.B. Soares Lab
cDNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward.

```

```

FEATURES             Location/Qualifiers
     source            1..516
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="IMAGE:3053711"
                     /tissue_type="lymph"
                     /cell_type="germinal center B cells"
                     /cell_line="MGC85"
                     /lab_host="DH10B (LTI)"
                     /note="Vector: pT73-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(0.5-1.5kb). Directionally cloned. Cells provided by Louis
M. Staudt, Ph.D. Library preparation by Maria de Fatima
Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT      112 a   132 c   150 g   122 t
ORIGIN

Query Match      32.6%; Score 182; DB 9; Length 516;
Best Local Similarity 70.3%; Pred. No. 2.1e-43;
Matches 277; Conservative 0; Mismatches 105; Indels 12; Gaps 2;

QY 10 gacaggtgcagctgcagagtcagagggagagattggtgcagggctggggctctctgcagac 69
DB 129 GTCAGGTGCAGCTGGTGGAGTCTGGGGGAGGCGGTGGTCCAGCCTGGGAGGTCCCTGAGAC 188
QY 70 tctcctgtgcagcctcgggacgcgcaccagtggtcactatggtcactatggtatggctggt 129
DB 189 TCTCCTGTGACGCTCTGGATTCACTTCACTAGTAG-----CTATGCTATGCACTGGG 239
QY 130 tccgccaggttccaggaagagcgtgagttgtgcagcattagtggtgagtggtgaaag 189
DB 240 TCCGCCAGGCTCCAGGCAAGGGGCTGGAGTGGGTGGCAGTTATATCATATGATGGAAGTA 299
QY 190 agacatggtataaagactccgtgaagggcgattccaccatctccagagataacgccaaaga 249
DB 300 ATAAATACTACGACAGCTCCCTGTAAGGGCCGATTCCACCATCTCCAGAGACAATTCACAAGA 359
QY 250 ctacggtttattctcaaatgaacagcctgaacctgaagatacgcgcgtttattattatg 309
DB 360 ACAGCTGTATCTGCAAAATGAACAGCCTGAGAACTGAGGACACGCGCTGTGTATTACTGTG 419
QY 310 ccgctcgaccggtcc---gcgtgagatgatatttccctccggttgggtttgactactggg 366
DB 420 CGAAGGCCCTCGTCTACGCGTGACTTCACTACTACTACGCTATGCGGCTCTGGG 479
QY 367 gccaggggacccaggtcacccgtctctctccaggatc 400
DB 480 GCCAAGGGACACGCTCACCCTCTTTTCCAGCTTC 513

Search completed: August 13, 2002, 11:59:15
Job time: 18868 sec

```

OM of: US-09-742-690-1 to: SPTREMBL_19:* out_format : pfs
 Date: Aug 13, 2002 1:53 PM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
 -Q=/cgn2_1/USPRO_spool/US09742690/runat_12082002_094915_166/app_query.fasta_1.623
 -DB=SPTREMBL_19 -QFMT=fastan -SUFFIX=rspt -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000
 -GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
 -USER=US09742690_@CGN1_1_132 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-742-690-1
 Query length: 558
 Database: SPTREMBL_19:*
 Database sequences: 562222
 Database length: 172994929
 Search time (sec): 101.840000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Document	hy
sp_human:Q96BB9	+ 402.50	698.17	2.6e-31	597	! Q96bb9 homo sapiens	(human).
sp_human:Q9UL90	+ 389.50	690.98	3.4e-30	113	! Q9ul90 homo sapiens	(human).
sp_human:Q9UL90	+ 379.50	672.65	7.4e-29	119	! Q920e7 mus musculus	(mouse).
sp_human:Q9HCC1	+ 376.00	666.99	7.5e-29	112	! Q9hcc1 homo sapiens	(human).
sp_human:Q9UL93	+ 369.50	655.06	3.3e-28	116	! Q9ul93 homo sapiens	(human).
sp_human:Q9UL71	+ 369.50	654.65	3.4e-28	121	! Q9ul71 homo sapiens	(human).
sp_human:Q99KA4	+ 369.00	640.38	5.2e-28	487	! Q99ka4 mus musculus	(mouse).
sp_human:Q9Y509	+ 368.00	650.11	4.9e-28	147	! Q9y509 homo sapiens	(human).
sp_human:Q9UL72	+ 365.00	646.87	9.3e-28	118	! Q9ul72 homo sapiens	(human).
sp_human:Q96K68	+ 364.00	631.32	1.6e-27	494	! Q96k68 homo sapiens	(human).
sp_human:Q9UL91	+ 362.00	641.52	1.9e-27	118	! Q9ul91 homo sapiens	(human).
sp_human:Q9UL84	+ 360.00	637.63	3.0e-27	122	! Q9ul84 homo sapiens	(human).
sp_human:Q9LWP5	+ 353.50	612.89	1.8e-26	479	! Q9lwp5 mus musculus	(mouse).
sp_human:Q9UL207	+ 346.50	600.27	8.9e-26	486	! Q9ul207 mus musculus	(mouse).
sp_human:Q9UL205	+ 342.00	592.50	2.5e-25	473	! Q9ul205 mus musculus	(mouse).
sp_human:Q9ULB6	+ 338.50	601.69	3.8e-25	95	! Q9ulb6 homo sapiens	(human).
sp_human:Q9UL88	+ 327.50	578.98	5.1e-24	131	! Q9ul88 homo sapiens	(human).
sp_human:Q9QYF0	+ 327.50	571.08	6.2e-24	298	! Q9qyf0 mus musculus	(mouse).
sp_human:Q9LXEL	+ 325.50	562.93	1.1e-23	480	! Q9lxel mus musculus	(mouse).
sp_human:Q9R1A4	+ 305.00	527.27	1.2e-21	437	! Q9r1a4 mus musculus	(mouse).
sp_mammal:Q9N0W6	+ 295.00	521.54	8.5e-21	124	! Q9n0w6 oryctolagus cuniculus	(r
sp_mammal:Q9N0W4	+ 293.00	517.98	1.3e-20	124	! Q9n0w4 oryctolagus cuniculus	(r
sp_human:Q9UGP3	+ 289.50	512.71	2.9e-20	112	! Q9ugp3 homo sapiens	(human).
sp_human:Q99NG4	+ 288.00	503.94	8.3e-20	121	! Q99ng4 mus musculus	(mouse).
sp_human:Q9LX92	+ 280.00	481.74	3.6e-19	482	! Q9lx92 mus musculus	(mouse).
sp_human:Q9UL87	+ 279.50	495.59	2.8e-19	104	! Q9ul87 homo sapiens	(human).
sp_human:Q9LX16	+ 279.00	486.62	3.9e-19	241	! Q9lxl6 mus musculus	(mouse).
sp_human:Q9UL92	+ 276.00	487.66	6.6e-19	124	! Q9ul92 homo sapiens	(human).
sp_human:Q9UL75	+ 275.00	486.03	8.2e-19	122	! Q9ul75 homo sapiens	(human).
sp_human:Q96KX8	+ 268.00	460.06	5.6e-18	496	! Q96kx8 homo sapiens	(human).
sp_human:Q95973	+ 267.00	469.77	5.4e-18	150	! Q95973 homo sapiens	(human).
sp_human:Q96EY0	+ 264.00	450.89	1.5e-17	613	! Q96ey0 homo sapiens	(human).
sp_human:Q924Q2	+ 262.50	462.28	1.5e-17	142	! Q924q2 mus musculus	(mouse).
sp_human:Q99LC4	+ 261.00	448.24	2.8e-17	463	! Q99lc4 mus musculus	(mouse).
sp_human:Q924R4	+ 259.00	455.83	3.3e-17	145	! Q924r4 mus musculus	(mouse).
sp_human:Q924R1	+ 258.00	454.05	4.2e-17	145	! Q924r1 mus musculus	(mouse).
sp_human:Q924P7	+ 258.00	454.05	4.2e-17	145	! Q924p7 mus musculus	(mouse).
sp_human:Q96Q80	+ 258.00	453.16	4.3e-17	159	! Q96qs0 homo sapiens	(human).
sp_human:Q924Q1	+ 257.50	453.36	4.7e-17	142	! Q924q1 mus musculus	(mouse).
sp_human:Q921C4	+ 257.00	454.25	5.0e-17	118	! Q921c4 mus musculus	(mouse).

sp_rodent:Q924Q8 + 256.50 451.31 5.9e-17 146 ! Q924q8 mus musculus (mouse).
 sp_rodent:Q924Q3 + 256.50 451.31 5.9e-17 146 ! Q924q3 mus musculus (mouse).
 sp_rodent:Q9QXF0 + 255.50 451.65 7.0e-17 117 ! Q9qxf0 mus musculus (mouse).
 sp_human:Q9UL73 + 255.50 451.49 7.1e-17 119 ! Q9ul73 homo sapiens (human).
 sp_rodent:Q99LA6 + 255.50 438.01 9.8e-17 484 ! Q99la6 mus musculus (mouse).
 seq_name: sp_human:Q96BB9
 seq_documentation_block:
 ID Q96BB9 PRELIMINARY; PRT; 597 AA.
 AC Q96BB9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL 65.0 KDA PROTEIN.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP TISSUE=PRIMARY B-CELLS FROM TONSILS;
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC015760; AAH15760.1;
 KW Hypothetical protein.
 SQ SEQUENCE 597 AA; 4FCA3AD8ECE263D9 CRC64;

alignment_scores:

Quality: 402.50 Length: 130
 Ratio: 3.946 Gaps: 1
 Percent Similarity: 78.462 Percent Identity: 65.385

alignment_block:

US-09-742-690-1 x Q96BB9
 Align seg 1/1 to: Q96BB9 from: 1 to: 597
 12 CAGGTGCAGCTCCAGGAGCTCAGGGGAGGATGTCAGGCTGGGGGCTC 61
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 20 GValGlnLeuLeuGlySerGlyGlyGlyLeuValGlnProGlyGlyse 36
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 62 TCTGAGACTCTCTGTCAGCTCCGAGCTCCGGGACGCCACCATGTCATGTC 111
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 36 rLeuArgLeuSerCysAlaAlaSerGlyPheSerPheSer..... 50
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 112 ACTATGTCATGGCTGGTTCGCCAGGTTCCAGGAGGAGCGTGAGTTT 161
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 51 ..TyrAlaMetAsnTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 162 GTCGCGAGCTATTAGTCGAGTCGTAAGAGACATGCTATAAGACTCCGT 211
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 67 ValSerAlaIleSerGlySerGlySerGlySerThrThrThrAlaAspSerVa 83
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 212 GAAGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 83 llysGlyArgPheThrIleSerArgAspAsnSerArgAspThrLeuTyrL 100
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 262 TCAAAATGAACAGCTGAAACCTCAAGATACGCCGCTTTATTATTGTC 311
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 116
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 312 GCTGACCGGTCGCGTCGAGTCATGATTTCCCTCCGCGGTGGTGGTACTA 361
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 117 LysAspProArgGlyTyrSerAlaSerGlyAsnTyrThrArgGluAspTy 133
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 362 CTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 401
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 133 rtrpGlyGlyGlyThrLeuValThrValSerSerGlySer 146
 :::::::::::::: :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
 seq_name: sp_human:Q9UL90

```

seq_name: sp_rodent:Q920E7

seq_documentation_block:
ID Q920E7 PRELIMINARY; PRT; 119 AA.
AC Q920E7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PTERIN-MIMICKING ANTI-IDIOPE HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_taxid=10090;
RX [1]
RP SEQUENCE FROM N.A.
RA Atkin J.D., Iape A., Jennings I.G., Horaitis O., Cotton R.G.H.;
RT "Definition of the Idiotope of Pterin-Mimicking Antibodies Expressed
in Mammalian Cells";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF307937; AAL09421.1; -
FT NON_TER 1
FT NON_TER 119
FT NON_TER 119
SQ SEQUENCE 119 AA; 13025 MW; F6E904044381CA7C CRC64;

alignment_scores:
Quality: 379.50 Length: 128
Ratio: 3.757 Gaps: 3
Percent Similarity: 78.906 Percent Identity: 64.062

alignment_block:
US-09-742-690-1 x Q920E7 ..
Align seg 1/1 to: Q920E7 from: 1 to: 119

12 CAGGTCCAGCTGCAGAGTCAGGGGAGGATGGTCAGCGCTGGGGGCTC 61
1 GluValGlnLeuValGluSerGlyGlyAspLeuValValProGlyGly 17
62 TCTGAGACTCTCTGTGCAGCCTCGGGAGCGGCCACAGTCGTGTCATGGTC 111
17 rLeuLysLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
112 ACTATGTTATGGCTGGTTCGCGCCAGGTTCCAGGGAAGAGCGTAGTTT 161
32 ..TyrGlyMetSerTrpValArgGlnThrProAspLysArgLeuGluTrp 47
162 GTCGCGAGCTATTAGTGGAGTGGTGGTAAGACACATGGTATAAAGACTCCGT 211
48 ValAlaThrIleSerSerGlyGlySerTyrThrTyrTrpProAspSerVa 64
212 GAAGGGCCGATTACCATCTCCAGAGATAAACGCAAGACTACGTTTATC 261
64 llysGlyArgPheThrIleSerArgAspAsnAlaLysAsnThrLeuTyrL 81
262 TGCAAATGAACAGCCCTGAACCTGAAGATACGGCCGCTTATTATTGTGCC 311
81 euGlnMetSerSerLeuLysSerGluAspThrAlaMetTyrTrpCysAla 97
312 GCTCGACCGGTCCGCGTGGATGATATTTCCCTCCGCGTGGGTTTGACTA 361
98 .....ArgHisGlyAspTyrAsp.....ValGlyPheAlaTy 108
362 CTGGGGCCAGGGGACCCAGGTACCGCTCCTCCITCA 395
108 rTrpGlyGlnGlyThrLeuValThrValSerAla 119

seq_name: sp_human:Q9HCC1

seq_documentation_block:
ID Q9HCC1 PRELIMINARY; PRT; 112 AA.
AC Q9HCC1;

```

```

DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SINGLE CHAIN FV (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX SEQUENCE FROM N.A.
RA Kikuchi M., Takeda C., Tsujimoto Y., Asada S., Nagata K.;
RT "An antibody fragment2A3 specific for native lysozyme :Isolation from a
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049915; BAB16829.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_Like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 1.
FT NON_TER 1
FT NON_TER 112
SQ SEQUENCE 112 AA; 12243 MW; 24FLA45EC3B84788 CRC64;

alignment_scores:
  Quality: 376.00      Length: 124
  Ratio: 3.958        Gaps: 2
  Percent Similarity: 76.613      Percent Identity: 61.290

alignment_block:
US-09-742-690-1 x Q9HCC1 ..
Align seg 1/1 to: Q9HCC1 from: 1 to: 112

12 CAGGTGCAGCTCAGAGTCAGGGGAGGATTGGTCAGCGTGGGGGCTC 61
   ::::::::::::::::::::
1 GluValGlnLeuValGluSerGlyGlyGlyValValArgProGlyGlySe 17
62 TCTGAGACTCTCTGTGCGAGCTCGGAGCGGCCGCCACGCTGCTGTC 111
   ::::::::::::::::::::
17 rLeuArgIleSerCysAlaAlaSerGlyPheThr.....PheAspA 31
112 ACTATGTGATGGCTGCTTCGCCAGGTTCAGGGAAGAGCGCTGAGTT 161
   ::::::::::::::::::::
31 spTyrglyMetSerTrpValArgGlnAlaProGlyGlyGlyLeuGluTrp 47
162 GTCCGACGATTATTAGTGGAGTGGTAAAGACATGCTATAAGACTCCGT 211
   ::::::::::::::::::::
48 ValSerGlyIleAsnTrpAsnGlyGlySerThrGlyTyraAlaAspSerVa 64
212 GAAGGCCGATTCCACCTCTCCAGAGATACGCCACACTACGGTTATC 261
   ::::::::::::::::::::
64 llysGlyArgPheThrIleSerArgAspAsnAlaLysAsnSerLeuTyrl 81
262 TGCAAATGAACACCTGAAACCTGAAGATACGGCCGTTTATTATTGTGCC 311
   ::::::::::::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrrCysAla 97
312 GCTCGACCGGTCCGCGTGGATGATATTTCCTCCCGGTGGGTTGACTA 361
   ::::::::::::::::::::
98 ArgArg.....ArgTyraAlaLeuAspTy 105
362 CTGGGCCGAGGACCCAGGTC 383
105 rTrpGlyGlnGlyThrLeuVal 112

seq_name: sp_human:Q9UL93

```

```

seq_documentation_block:
ID Q9UL93 PRELIMINARY; PRT; 116 AA.
AC Q9UL93;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX SEQUENCE FROM N.A.
RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035021; AAD56257.1; -.
DR HSSP; P01772; 2FB4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 116
SQ SEQUENCE 116 AA; 12434 MW; 0DA0348154DD6061 CRC64;

alignment_scores:
  Quality: 369.50      Length: 127
  Ratio: 3.732        Gaps: 2
  Percent Similarity: 77.953      Percent Identity: 61.417

alignment_block:
US-09-742-690-1 x Q9UL93 ..
Align seg 1/1 to: Q9UL93 from: 1 to: 116

15 GTCAGCTCAGGAGTCAGGGGAGGATTGGTCAGCGTGGGGGCTCTCT 64
   ::::::::::::::::::::
1 ValGlnLeuValGluSerGlyGlyGlyValValGlnProGlyArgSerLe 17
65 GAGACTCTCTGTCAGCTCGGAGCGGCCGCCACCGCTGCTCATGCTCACT 114
   ::::::::::::::::::::
17 uArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer.....T 31
115 ATGATGCGCTGCTTCGCCAGGTTCAGGGAAGAGCGCTGAGTTTGTCT 164
   ::::::::::::::::::::
31 yAlaMetHisTrpValArgGlnAlaProGlyGlyGlyLeuGluTrpVal 47
165 GCAGCTATTAGTGGAGTGGTAAAGACATGCTATAAGACTCCGTGAA 214
   ::::::::::::::::::::
48 AlaValIleSerTyraSpGlySerAsnLysTyrrTyraAlaAspSerVally 64
215 GGGCCGATTCCACATCTCCAGAGATACGCCACACTACGGTTTATCTGC 264
   ::::::::::::::::::::
64 sGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrrLeuG 81
265 AAATGAACACCTGAAACCTGAAGATACGGCCGTTTATTATTGTGCCGT 314
   ::::::::::::::::::::
81 InMetAsnSerLeuArgAlaGluAspThrAlaMetTyrrTyrrCysAlaGly 97
315 CGACCGCTCGCGTGGATGATATTTCCTCCCGGTGGGTTGACTACTG 364
   ::::::::::::::::::::
98 .....GlyGlyGlyLeuGlyLeuGlyTyrr 106
365 GGGCCGAGGAGCCAGGTCACCGTCTCTCTCA 395
   ::::::::::::::::::::
106 pGlyGlnGlyThrLeuValThrValSerSer 116

```


354TTTGACTACTGGGCCAGGGCCAGGCCAGGTCACCGTCTCTCAGGAT 399
 |||||
 128 rArgPheAspTyrTrpGlyGlnGlyThrThrIleThrValSerSerGln 145
 400 CT 401
 ||
 145 er 145

seq_name: sp_human:Q9Y509

seq_documentation_block:
 ID Q9Y509 PRELIMINARY; PRT; 147 AA.
 AC Q9Y509;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE V3 PROTEIN (FRAGMENT).
 GN V3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96071149; PubMed=7475288;
 RA Cao J., Vescio R.A., Rettig M.B., Hong C.H., Kim A., Lee J.C.,
 Lichtenstein A.K., Berenson J.R.;
 RT "A CD10-positive subset of malignant cells is identified in multiple
 myeloma using PCR with patient-specific immunoglobulin gene primers.";
 RL Leukemia 9:1948-1953(1995).
 DR EMBL: S80860; AAD14339.1;
 DR HSSP: P01772; 2F84.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_v.
 DR Pfam: PF00047; Ig_1.
 DR SMART: SM00406; IGV; 1.
 FT NON_TER 1
 SQ SEQUENCE 147 AA; 15768 MW; 8489FCNA7BC925C CRC64;

alignment_scores:
 Quality: 368.00 Length: 131
 Ratio: 3.644 Gaps: 2
 Percent Similarity: 77.099 Percent Identity: 60.305

alignment_block:

US-09-742-690-1 x Q9Y509 ..

Align seg 1/1 to: Q9Y509 from: 1 to: 147

12 CAGGTGCGAGTCGAGGAGTCAGGGGAGGATGTGTCAGGCTGGGGCTC 61
 |||||
 1 GlnValHisLeuValGluSerGlyGlyValValGlnProGlyLys 17
 62 TCTCAGACTCTCTGTCAGCCTCGGACCGCCAGCCAGCTGGTCATGTC 111
 |||||
 17 rLeuArgLeuSerCysGlnAlaSerGlyPheThrPheSer.....T 31
 112 ACTATGTTATGGCTGGTTCGCGCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
 |||||
 31 hrTyrGlyMetSerTyrValArgGlnAlaProGlyLysGlyLeuAspTrp 47
 162 GTCGCACTATTAGGTGGAGTGGTAAGACATGGTATTAAGACTCCGT 211
 |||||
 48 ValAlaLeuIleSerTyrAspGlySerThrGlnTyrTyrAlaGlySerVa 64
 212 GAAGGGCCGATTACCATCTCCAGATACGCCAAGACTACGTTTATC 261
 |||||
 64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
 262 TGCAATGAACAGCCTCAACCTCAACATAGCGCGTTTATTATGTCGCC 311
 |||||

81 euGlnMetThrSerLeuArgValCluAspThrAlaValTyrTyrCysAla 97
 312 GCTCGACCGTCCGGCGGAGATATTTCC...CTGCCGGTTGGTTGA 358
 |||||
 98 LysAspGlyAsnTyrPheAspSerValGlyTyrTyrTyrAlaGlyIleAs 114
 359 CTACTGGGGCCAGGGACCCAGGTCACCGTCTCTCAGGATCT 401
 |||||
 114 pTyrTrpGlyGlnGlyThrLeuValThrValSerSerAlaSer 128
 seq_name: sp_human:Q9UL72

seq_documentation_block:
 ID Q9UL72 PRELIMINARY; PRT; 118 AA.
 AC Q9UL72;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
 DE (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98277139; PubMed=9614934;
 RA Wu X., Liu B., Van der Merwe P.L., Kalis N.N., Berney S.M.,
 Young D.C.;
 RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
 fetus.";
 RL Clin. Immunol. Immunopathol. 87:184-192(1998).
 DR EMBL: AF035042; RAD56278.1;
 DR HSSP: P01772; 2F84.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003596; Ig_v.
 DR Pfam: PF00047; Ig_1.
 DR SMART: SM00406; IGV; 1.
 FT NON_TER 1
 SQ SEQUENCE 118 AA; 12872 MW; B4D1A5944B2D5CCA CRC64;

alignment_scores:

Quality: 365.00 Length: 128
 Ratio: 3.763 Gaps: 3
 Percent Similarity: 75.781 Percent Identity: 62.500

alignment_block:

US-09-742-690-1 x Q9UL72 ..

Align seg 1/1 to: Q9UL72 from: 1 to: 118

12 CAGGTGCGAGTCGAGGAGTCAGGGGAGGATGTGTCAGGCTGGGGCTC 61
 |||||
 1 GluValGlnLeuValGluSerGlyGlyLeuValGlnProGlyLys 17
 62 TCTCAGACTCTCTGTCAGCCTCGGACCGCCAGCTGGTCATGTC 111
 |||||
 17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrValSerSerAsn.... 32
 112 ACTATGTTATGGCTGGTTCGCGCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
 |||||
 33TyrMetAsnTrpValArgGlnAlaProGlyLysGlyLeuGluSer 47
 162 GTCGCACTATTAGGTGGAGTGGTAAGACATGGTATTAAGACTCCGT 211
 |||||
 48 Val...SerValThrTyrSerGlySerSerTyrTyrAlaAspSerVa 63
 212 GAAGGGCCGATTACCATCTCCAGATACGCCAAGACTACGTTTATC 261
 |||||
 63 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 80


```

262 TGCAATGAACAGCTGAACCTGAAGATACGCGCTTATTATTGTGCC 311
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrCysAla 97
312 GCTGACCGGTCGCGGTGGATGATATTTCCCTGCGCGTGGTGGACTA 361
98 .....ArgGlyAspSerGlu.....AlaPheAspIrl 107
362 CVGGGGCCAGGGACCCAGGTACCGTCTCC 392
107 etrpGlyGlnGlyThrMetValThrValSer 117

seq_name: sp_human:Q9UL84

seq_documentation_block:
ID Q9UL84 PRELIMINARY; PRT; 122 AA.
AC Q9UL84;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MYOSIN-REACTIVE IMMUNOGLOBULIN HEAVY CHAIN VARIABLE REGION
DE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98277139; PubMed=9614934;
RA Wu X., Liu B., Van der Merwe P.L., Kallis N.N., Berney S.M.,
RA Young D.C.;
RT "Myosin-reactive autoantibodies in rheumatic carditis and normal
RT fetus.";
RL Clin. Immunol. Immunopathol. 87:184-192(1998).
DR EMBL; AF035030; AAD56266.1; -
DR HSP; P01772; 2FBA.
DR InterPro; IPR003006; Iq_MHC.
DR InterPro; IPR003596; Iq_v.
DR Pfam; PF00047; Iq; 1.
DR SMART; SM00406; IGV; 1.
FT NON_TER 1
FT NON_TER 122
SQ SEQUENCE 122 AA; 13579 MW; 36054D41366545B8 CRC64;

```

```

alignment_scores:
  Quality: 360.00      Length: 130
  Ratio: 3.673        Gaps: 4
  Percent Similarity: 75.385  Percent Identity: 63.077

alignment_block:
US-09-742-690-1 x Q9UL84 ..
Align seg 1/1 to: Q9UL84 from: 1 to: 122

12 CAGGTGCAGCTCAGAGTCAGGGAGGAGTGGTCAGGCTGGGGGCTC 61
1 GluValGlnLeuValGluSerGlyGlyValValGlnProGlyArgSe 17
62 TCTGAGACTCTCTGTGACGCTCGGAGCGCCACCGTGGTCATGGTC 111
17 rLeuArgLeuSerCysAlaAlaSer....ArgPheThr.....PheSerA 31
112 ACTATGGTATGGCTGTCCGCCAGGTTCAGGGAAGGAGCGGTGATTT 161
31 snfyGlyMethisThrValArgGlnAlaProGlyLysGlyLeuGluTrp 47
162 GTCGACGCTATTAGTGGAGTGTAAAGACATGCTATAAAGACTCCGT 211
48 ValAlaAlaIleSerAsnAspGlySerAsnLysPheTyrAlaAspSerVa 64
212 GAAGGGCCGATTCCACCTCCAGAGATAACGCCAAGACTACGGTTATC 261

```

```

64 llyGlyArgPheThrIlePheArgAspAsnSerLysAsnMetMetAspL 81
262 TGCAATGAACAGCTGAACCTGAAGATACGCGCTTATTATTGTGCC 311
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrCysAla 97
312 GCTGACCGGTCGCGGTGGATGATATTTCCCTGCGCGTGGG.....TT 355
98 Lys.....AspGluArgGlyArgLeuValGlyThrTyrPh 109
356 TGACTACTGGGCGCCAGGGACCCAGGTACCGTCTCTCTCA 395
109 easPtyrTrpGlyGlnGlyThrLeuValThrValSer 122

seq_name: sp_rodent:Q91WP5

seq_documentation_block:
ID Q91WP5 PRELIMINARY; PRT; 479 AA.
AC Q91WP5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 51.6 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC013656; AAH13656.1; -
KW Hypothetical protein.
SQ SEQUENCE 479 AA; 51603 MW; ECB2D087748584F CRC64;

alignment_scores:
  Quality: 353.50      Length: 128
  Ratio: 3.682        Gaps: 2
  Percent Similarity: 75.000  Percent Identity: 57.812

alignment_block:
US-09-742-690-1 x Q91WP5 ..
Align seg 1/1 to: Q91WP5 from: 1 to: 479

12 CAGGTGCAGCTCAGAGTCAGGGAGGAGTGGTCAGGCTGGGGGCTC 61
20 GluValGlnLeuValGluSerGlyGlyGlyLeuValLysProGlyGly 36
62 TCTGAGACTCTCTGTGACGCTCGGAGCGCCACCGTGGTCATGGTC 111
36 rLeuLysValSerCysAlaAlaSerGlyLeuThrPheSer.....A 50
112 ACTATGGTATGGCTGTCCGCCAGGTTCAGGGAAGGAGCGGTGATTT 161
50 snfyAlaMetSerTrpValArgGlnSerProGlyLysArgLeuGluTrp 66
162 GTCGACGCTATTAGTGGAGTGTAAAGACATGCTATAAAGACTCCGT 211
67 ValAlaAlaIleAsnSerAsnGlyGlyAsnThrTyrTrpSerAspThr 83
212 GAAGGGCCGATTCCACCTCCAGAGATAACGCCAAGACTACGGTTATC 261
83 llyGlyArgPheThrIleSerArgAspAsnAlaLysSerThrLeuTrL 100
262 TGCAATGAACAGCTGAACCTGAAGATACGCGCTTATTATTGTGCC 311
100 euGlnMetSerSerLeuArgSerGluAspThrAlaPheTyrCysVal 116
312 GCTGACCGGTCGCGGTGGATGATATTTCCCTGCGCGTGGTGGACTA 361

```

Wed Aug 14 11:49:30 2002

```
117 ArgGlyGlyTyr.....PheAspVa 123
|||||
362 CTGGGGCCAGGGGACCCAGGTCACCGTCTCTCA 395
||||| ||||| ||||| ||||| |||||
123 lTrpGlyAlaGlyThrAlaValThrValSerSer 134
||||| ||||| ||||| ||||| |||||

seq_name: sp_rodent:Q91207
seq_documentation_block:
ID Q91207 PRELIMINARY; PRT; 486 AA.
AC Q91207;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 52.7 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010324; AAI10324.1; -.
KW Hypothetical protein.
SQ SEQUENCE 486 AA; 52682 MW; 4FEF835125DA870B CRC64;

alignment_scores:
Quality: 346.50 Length: 136
Ratio: 3.397 Gaps: 5
Percent Similarity: 75.000 Percent Identity: 55.882

alignment_block:
US-09-742-690-1 x Q91207 ..
Align seg 1/1 to: Q91207 from: 1 to: 486

12 CAGGTGCAGCTGCAGGAGTCCAGGGGAGGATGGTGACAGGCTGGGGGCTC 61
:::||||| ||||| ||||| ||||| |||||
20 GluValHisLeuValGluSerGlyGlyLeuValLysProGlyGly 36
||||| ||||| ||||| ||||| |||||
62 TCTGAGACTCTCTGTGAGCTCGGAGCGCCAGCCAGTGGTCATGTC 111
||||| ||||| ||||| ||||| |||||
36 rLeuLysLeuSerCysValValSerGlyPheSerPheThrSer..... 50
||||| ||||| ||||| ||||| |||||
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGAGGAGCGTGAGTTT 161
||| ||||| ||||| ||||| |||||
51 ..TyrAspMetSerTrpValArgGlnThrProGluArgArgLeuGluTrp 66
||||| ||||| ||||| ||||| |||||
162 GTCGACGCTATTAGTGAGTGGTAAAGACATGGTATTAAGACTCCGT 211
||||| ||||| ||||| ||||| |||||
67 ValAlaAlaIleThr...SerGlyGlyAsnThrTyrTyrProAspAsnVa 82
||||| ||||| ||||| ||||| |||||
212 GAAGGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTATC 261
||||| ||||| ||||| ||||| |||||
82 lIysGlyArgPheThrValSerArgAspAsnAlaLysTyrThrLeuTyrL 99
||||| ||||| ||||| ||||| |||||
262 TGCAATGAACAGCTGAACCTGAAGATACGGCGTTTATTATTGTGCC 311
||||| ||||| ||||| ||||| |||||
99 euGlnMetSerSerLeuLysSerGluAspThrAlaMetTyrTyrCys... 114
||||| ||||| ||||| ||||| |||||
312 GCTCAGACCGTCCGCGTGAGATGATATTCCTCGCGGTT..... 350
||||| ||||| ||||| ||||| |||||
115 ValArgPro.....GluIleProIleTyrTyrSer 125
||||| ||||| ||||| ||||| |||||
351 .....GGGTTTGACTACTGGGGGACCCAGGTCACCGTCTCCT 393
||||| ||||| ||||| ||||| |||||
125 rGlySerTyrPheAspSerTrpGlyGlnGlyThrThrIleThrValSer 142
||||| ||||| ||||| ||||| |||||
394 CAGGATCT 401
|| |||
```

```
142 erGluSer 144
seq_name: sp_rodent:Q91205
seq_documentation_block:
ID Q91205 PRELIMINARY; PRT; 473 AA.
AC Q91205;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 51.9 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010327; AAI10327.1; -.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;

alignment_scores:
Quality: 342.00 Length: 128
Ratio: 3.600 Gaps: 3
Percent Similarity: 74.219 Percent Identity: 58.594

alignment_block:
US-09-742-690-1 x Q91205 ..
Align seg 1/1 to: Q91205 from: 1 to: 473

12 CAGGTGCAGCTGCAGGAGTCCAGGGGAGGATGGTGACAGGCTGGGGGCTC 61
:::||||| ||||| ||||| ||||| |||||
20 GluValGlnLeuValGluSerGlyGlyLeuValLysProGlyGly 36
||||| ||||| ||||| ||||| |||||
62 TCTGAGACTCTCTGTGAGCTCGGAGCGCCAGCCAGTGGTCATGTC 111
||||| ||||| ||||| ||||| |||||
36 rArgLysLeuSerCysAlaAlaSerGlyPheThrPheSer.....A 50
||||| ||||| ||||| ||||| |||||
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGAGGAGCGTGAGTTT 161
||||| ||||| ||||| ||||| |||||
50 sPTyrGlyMetHisTrpValArgGlnAlaProGluLysGlyLeuGluTrp 66
||||| ||||| ||||| ||||| |||||
162 GTCGACGCTATTAGTGAGTGGTAAAGACATGGTATTAAGACTCCGT 211
||||| ||||| ||||| ||||| |||||
67 ValAlaTyrIleAsnSerGlySerThrThrIleTyrTyrAlaAspThrVa 83
||||| ||||| ||||| ||||| |||||
212 GAAGGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTATC 261
||||| ||||| ||||| ||||| |||||
83 lLysGlyArgPheThrIleSerArgAspAsnAlaLysAsnThrLeuPheL 100
||||| ||||| ||||| ||||| |||||
262 TGCAATGAACAGCTGAACCTGAAGATACGGCGTTTATTATTGTGCC 311
||||| ||||| ||||| ||||| |||||
100 euGlnMetThrSerLeuArgSerGluAspThrAlaMetTyrTyrCys... 115
||||| ||||| ||||| ||||| |||||
312 GCTCAGACCGTCCGCGTGAGATGATATTCCTCGCGGTTGGGTTTGACTA 361
||||| ||||| ||||| ||||| |||||
116 AlaArgGluLeuTrpLeuArgArgIle.....AspTyr 126
||||| ||||| ||||| ||||| |||||
362 CTGGGGCGGAGGGGACCCAGGTCACCGTCTCTCTCA 395
||||| ||||| ||||| ||||| |||||
126 rTrpGlyGlnGlyThrThrIleThrValSerSer 137
||||| ||||| ||||| ||||| |||||
```

OM of: US-09-742-690-1 to: SwissProt_40:* out_format : pfs

Date: Aug 13, 2002 1:54 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+np.model -DEV=xlp
-Q/cgn2_1/USPTO_spoal/US09742690/runat_12082002_094916_213/app_query.fasta_1.623
-DB=SwissProt_40 -QPMT=fastan -SUFFIX=rsp -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DEPOP=6.000 -DEPEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MTN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZ=500 -MTNLEN=0 -MAXLEN=2000000000
-USER=US09742690 -CIGN1_1_28 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-742-690-1
Query length: 558
Database: SwissProt_40:*
Database sequences: 105224
Database length: 38719550
Search time (sec): 40.820000

score_list:	Sequence	Strd Orig	ZScore	EScore	Len	Documentation
SwissProt_40:HV3J_HUMAN	+	386.50	634.02	1.1e-27	121	P01771 homo sapiens (human)
SwissProt_40:HV3G_HUMAN	+	369.00	605.24	4.1e-26	122	P01768 homo sapiens (human)
SwissProt_40:HV3H_HUMAN	+	353.00	578.98	1.7e-24	122	P01769 homo sapiens (human)
SwissProt_40:HV3I_HUMAN	+	351.50	576.75	1.1e-24	119	P01770 homo sapiens (human)
SwissProt_40:HV3T_HUMAN	+	348.00	571.25	3.5e-24	116	P01781 homo sapiens (human)
SwissProt_40:HV3U_HUMAN	+	347.00	569.29	4.3e-24	117	P01782 homo sapiens (human)
SwissProt_40:HV3V_HUMAN	+	346.50	568.71	4.8e-24	120	P01764 homo sapiens (human)
SwissProt_40:HV3K_HUMAN	+	344.00	567.20	5.4e-24	126	P01772 homo sapiens (human)
SwissProt_40:HV3L_HUMAN	+	344.00	564.84	8.0e-24	114	P01784 canis familiaris (dog)
SwissProt_40:HV3F_HUMAN	+	343.50	563.94	8.9e-24	115	P01767 homo sapiens (human)
SwissProt_40:HV3E_HUMAN	+	343.50	563.94	9.0e-24	120	P01766 homo sapiens (human)
SwissProt_40:HV3A_HUMAN	+	341.00	559.30	1.5e-23	122	P01762 homo sapiens (human)
SwissProt_40:HV3J_MOUSE	+	340.50	559.02	1.7e-23	115	P01802 mus musculus (mouse)
SwissProt_40:HV3L_MOUSE	+	340.50	558.70	1.7e-23	119	P01773 homo sapiens (human)
SwissProt_40:HV32_MOUSE	+	337.50	554.10	3.2e-23	113	P01799 mus musculus (mouse)
SwissProt_40:HV30_MOUSE	+	336.50	552.62	3.9e-23	113	P01801 mus musculus (mouse)
SwissProt_40:HV31_MOUSE	+	333.50	547.70	7.3e-23	113	P01800 mus musculus (mouse)
SwissProt_40:HV02_CANFA	+	333.50	547.38	7.3e-23	117	P01785 canis familiaris (dog)
SwissProt_40:HV27_MOUSE	+	331.50	544.41	1.1e-22	113	P01796 mus musculus (mouse)
SwissProt_40:HV3D_HUMAN	+	331.50	544.25	1.1e-22	115	P01765 homo sapiens (human)
SwissProt_40:HV16_MOUSE	+	331.00	541.89	1.3e-22	136	P01783 mus musculus (mouse)
SwissProt_40:HV35_MOUSE	+	330.00	542.12	1.5e-22	111	P01804 mus musculus (mouse)
SwissProt_40:HV3B_HUMAN	+	329.00	540.23	1.9e-22	114	P01763 homo sapiens (human)
SwissProt_40:HV05_CARAU	+	328.00	538.43	2.3e-22	116	P19181 carassius auratus (goldfish)
SwissProt_40:HV55_MOUSE	+	326.50	535.89	3.2e-22	117	P18526 mus musculus (mouse)
SwissProt_40:HV28_MOUSE	+	325.50	534.57	3.9e-22	113	P01797 mus musculus (mouse)
SwissProt_40:HV29_MOUSE	+	321.50	528.01	9.1e-22	113	P01798 mus musculus (mouse)
SwissProt_40:HV53_MOUSE	+	319.50	524.41	1.4e-21	117	P18524 mus musculus (mouse)
SwissProt_40:HV41_MOUSE	+	314.50	516.20	4.0e-21	117	P01811 mus musculus (mouse)
SwissProt_40:HV57_MOUSE	+	313.50	516.19	4.8e-21	98	P18528 mus musculus (mouse)
SwissProt_40:HV54_MOUSE	+	313.50	514.56	4.9e-21	117	P18525 mus musculus (mouse)
SwissProt_40:HV52_MOUSE	+	313.50	514.10	5.0e-21	123	P01794 mus musculus (mouse)
SwissProt_40:HV3N_HUMAN	+	312.00	511.94	6.8e-21	119	P01775 homo sapiens (human)
SwissProt_40:HV20_MOUSE	+	312.00	511.71	6.8e-21	122	P01789 mus musculus (mouse)
SwissProt_40:HV22_MOUSE	+	311.50	510.82	7.6e-21	123	P01791 mus musculus (mouse)
SwissProt_40:HV3M_HUMAN	+	311.00	510.30	8.4e-21	119	P01774 homo sapiens (human)
SwissProt_40:HV19_MOUSE	+	308.00	507.27	1.5e-20	97	P18527 mus musculus (mouse)
SwissProt_40:HV16_MOUSE	+	307.50	504.26	1.8e-20	123	P01788 mus musculus (mouse)
SwissProt_40:HV21_MOUSE	+	307.00	503.51	2.0e-20	122	P01790 mus musculus (mouse)
SwissProt_40:HV39_MOUSE	+	306.50	503.00	2.2e-20	118	P01809 mus musculus (mouse)

SwissProt_40:HV01_RAT + 306.50 501.29 2.2e-20 142 i P01805 rattus norvegicus (r
SwissProt_40:HV38_MOUSE + 306.00 502.10 2.4e-20 119 i P01808 mus musculus (mous
SwissProt_40:HV3P_HUMAN + 306.00 502.10 2.4e-20 119 i P01777 homo sapiens (huma
SwissProt_40:HV18_MOUSE + 305.97 2.7e-20 123 i P01787 mus musculus (mous
SwissProt_40:HV34_MOUSE + 304.50 500.11 3.3e-20 113 i P01803 mus musculus (mous

seq_name: SwissProt_40:HV3J_HUMAN

seq_documentation_block:

ID HV3J_HUMAN STANDARD; PRT; 121 AA.
AC P01771;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-JII region HLL.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX MEDLINE=79124695; PubMed=420800;
RA Chiu Y.-Y.H., Lopez de Castro J.A., Poljak R.J.;
RT "Amino acid sequence of the VH region of human myeloma
cryoglobulin IgG HLL."
RL Biochemistry 18:553-560(1979).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC PROTEIN.
DR PIR: A02054; GIHULL.
DR HSSP; P01772; 2FE4.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam: PF00047; Ig: 1.
DR SMART; SM00406; IgV: 1.
KW Immunoglobulin V region.
FT MOD_RES 1 121
FT NON_TER 121 121
SQ SEQUENCE 121 AA; 13566 MW; 480FC53610EF5DAB CRC64;

alignment_scores:

Quality: 386.50 Length: 128
Ratio: 3.865 Gaps: 2
Percent Similarity: 78.125 Percent Identity: 61.719

alignment_block:

US-09-742-690-1 x HV3J_HUMAN ..
Align seg 1/1 to: HV3J_HUMAN from: 1 to: 121
12 CAGTGCAGCTGCAGCAGTCAGGCGGAGGATGTCAGGCTGGGGCTC 61
|||||:||||:||||:||||:||||:||||:||||:||||:||||:||||
1 GlnValLysLeuValGlnAlaGlyGlyValValGlnProGlyArgse 17
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||
62 TCTGAGACTCTCTGTGACGCTCGGCGGCGGCGGCGGCGGCGGCTC 111
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||
17 rLeuArgLeuSerCysIleAlaSerGlyPheThrPheSer.....A 31
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||
112 ACTATGATGATGGCTGGTTCGCCAGGTCAGGAGGAGGAGCGTGATTT 161
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||
31 snTyrGlyMetHisTyrValArgGlnAlaProGlyLysGlyLeuGluTrp 47
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||
162 GTCCGACCTATTAGGTGGAGTGGTAAAGACATGTTATTAAGACTCCGT 211
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||
48 ValAlaValIleTyrTrpAsnGlySerArgThrTyrTyrGlyAspSerVa 64
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||
212 GAAGGCGCGATTCACCATCTCCAGAGATAACGCAAGACTACGGTTATC 261
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||
64 LysGlyArgPheThrIleSerArgAspSerLysArgThrLeuTyrM 81
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||
262 TCGAAATGACACCGCTGAACCTCAAGATACGCGCGGCTTATTATGTC 311
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:||||
81 et***MetAsnSerLeuArgThrGluAspThrAlaValTyrTyrCysAla 97

312 GCTGACCGGTCCGGTGGATGATATTTCCCTGCGGTGGTGGTTGACATCA 361
 ||| ||||| :|||
 98 ArgAspPro.....AspIleLeuThrAlaPheSerPheAspTyr 110
 ||| ||||| :|||
 362 CTGGGGCCAGGACCCAGGTCCCTCCCTCA 395
 ||| ||||| :|||
 110 rTrpGlyGlnGlyValLeuValThrValSerSer 121

seq_name: SwissProt_40:HV3G_HUMAN

seq_documentation_block:

ID HV3G_HUMAN STANDARD; PRT; 122 AA.

AC P01768;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig heavy chain V-III region CAM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.

RX MEDLINE=81013859; PubMed=6774332;
 RA Lehman D.W., Putnam F.W.;
 RT "Amino acid sequence of the variable region of a human mu chain:
 RT location of a possible JH segment.";
 RL Proc. Natl. Acad. Sci. U.S.A. 77:3239-3243(1980).
 CC -1- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM THE PLASMA OF A
 CC PATIENT WITH MACROGLOBULINEMIA.

DR PIR: A02051; M3HUAM.
 DR HSSP; P01772; 21G2.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig_1.
 DR SMART; SM00406; IGv; 1.
 KW Immunoglobulin V region.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT NON_TER 122 122
 FT SEQUENCE 122 AA; 13668 MW; A42D0F17D252F1C2 CRC64;

alignment_scores:

Quality: 369.00 Length: 129
 Ratio: 3.690 Gaps: 3
 Percent Similarity: 77.519 Percent Identity: 59.690

alignment_block:

US-09-742-690-1 x HV3G_HUMAN ..

Align seg 1/1 to: HV3G_HUMAN from: 1 to: 122

12 CAGGTGACGTGTCAGGAGTCAGGGGAGGATTTGGTGCAGGCTGGGGCTC 61
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 1 GlnValGluLeuValGluSerGlyGlyValVal***ProGlyArgSe 17
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 62 TCTGACACTCTCTGTGACGCTCGGACGCCACACAGTGGTCATGGTC 111
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSer.....A 31
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 112 ACTATGATGGCTGTTCCCGCAGGTTCCAGGAGGAGGAGCTGAGTTT 161
 :|||:|||| |||||:|||||:|||||:|||||:|||||
 31 snfYrAlaMetHisTrpValArgGlnProProGlyLysGlyLeuGluTrp 47
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 162 GTCGACGCTATTAGTGGAGTGTGAAGACACATGGTATAAGACTCCCGT 211
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 48 ValAlaValIleSerTyr***Gly*****LysTyrTyrAla***SerVa 64
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 212 GAAGGCCGATTCACCATCTCCAGAGATACGCCAAGACTACGGTTTATC 261
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 64 lLysGlyArgPheThrIleSerArgAsp***SerLys***ThrLeuTyrL 81

262 TGCATATGACAGCCCTGAACCTGAAGATACGCGCTTTATTATTCTGCC 311
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 81 euGlnMetAsnSerLeuArgAlaGlu***ThrAlaValTyrTyrCysAla 97
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 312 GCT...CGACCGGTCCGGTGGATGATATTTCCCTGCGGTGGTGGTTGA 358
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 98 ArgAspArgProLeuTyrGly***TyrArgAla.....PheAs 110
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 359 CTACTGGGCGCCAGGACCCAGGTCCCTCCCTCA 395
 :|||:|||| |||||:|||||:|||||:|||||:|||||
 110 nTyrTrpGlyGlnGlyThrLeuValThrValSerSer 122

seq_name: SwissProt_40:HV3H_HUMAN

seq_documentation_block:

ID HV3H_HUMAN STANDARD; PRT; 122 AA.

AC P01769;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig heavy chain V-III region GA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=74175307; PubMed=4208843;
 RA Florent G., Lehman D., Putnam F.W.;
 RT "The switch point in mu heavy chains of human IgM immunoglobulins.";
 RL Biochemistry 13:2482-2498(1974).
 CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A WALDENSTROM'S
 CC MACROGLOBULIN.

DR PIR: A02052; M3HUGA.
 DR HSSP; P01772; 21G2.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig_1.
 DR SMART; SM00406; IGv; 1.
 KW Immunoglobulin V region.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT NON_TER 122 122
 FT SEQUENCE 122 AA; 13166 MW; 74E5B6959E84100A CRC64;

alignment_scores:

Quality: 353.00 Length: 128
 Ratio: 3.602 Gaps: 2
 Percent Similarity: 76.562 Percent Identity: 53.125

alignment_block:

US-09-742-690-1 x HV3H_HUMAN ..

Align seg 1/1 to: HV3H_HUMAN from: 1 to: 122

12 CAGGTGACGTGTCAGGAGTCAGGGGAGGATTTGGTGCAGGCTGGGGCTC 61
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 1 GlnVal***LeuVal***SerGlyGlyGlyAlaVal***ProGlyArgSe 17
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 62 TCTGACACTCTCTGTGACGCTCGGACGCCACACAGTGGTCATGGTC 111
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSer.....T 31
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 112 ACTATGATGGCTGTTCCCGCAGGTTCCAGGAGGAGGAGCTGAGTTT 161
 :|||:|||| |||||:|||||:|||||:|||||:|||||
 31 hrTyrAlaMetHisTrpValArgGlnAlaProGlyLysGlyLeu***Trp 47
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 162 GTCGACGCTATTAGTGGAGTGTGAAGACACATGGTATAAGACTCCCGT 211
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 48 LeuSerValIleSerTyr***Gly*****TyrTyrAlaAlaSerVa 64
 |||||:|||| |||||:|||||:|||||:|||||:|||||
 212 GAAGGCCGATTCACCATCTCCAGAGATACGCCAAGACTACGGTTTATC 261
 |||||:|||| |||||:|||||:|||||:|||||:|||||

```

64  LysGlyArgPheThrIleSerArg*****SerLys***ThrMetTyrL 81
262  TGCAATGACAGCCTGAACCTGAAGATACGGCGGTTTATTATTGTC 311
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
81  euGluMetAsnSerLeuArgAlaGluAsnThrAlaValTyrTyrCysAla 97
312  GCTGACCGGTCGCCGTGGATGATATTCCTCGCGGTGGTGGTTGACTA 361
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
98  ArgSerGlyIleAlaLeuGlySerValAla.....GlyThrAspTyr 111
362  CTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCA 395
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
111  rTrpGly***GlyThrLeuValThrIleSerSer 122
seq_name: SwissProt_40:HV3I_HUMAN
seq_documentation_block:
ID HV3I_HUMAN STANDARD; PRT; 119 AA.
AC P01770;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region NIE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
[1]
SEQUENCE.
RX MEDLINE=77070269; PubMed=826475;
RA Ponstingl H., Hilschmann N.;
RT "The rule of antibody structure. The primary structure of a
RT monoclonal IgG1 immunoglobulin (myeloma protein Nie). III. The
RT chymotryptic peptides of the H-chain, alignment of the tryptic
RT peptides and discussion of the complete structure."
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1571-1604(1976).
RN [2]
RP DISULFIDE BOND.
RX MEDLINE=77070267; PubMed=1002129;
RA Dreker L., Schwarz J., Reichel W., Hilschmann N.;
RT "Rule of antibody structure. The primary structure of a monoclonal
RT IgG1 immunoglobulin (myeloma protein Nie). I: Purification and
RT characterization of the protein, the L- and H-chains, the
RT cyanogen bromide cleavage products, and the disulfide bridges."
RL Hoppe-Seyler's Z. Physiol. Chem. 357:1515-1540(1976).
CC -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGG1 MYELOMA
CC PROTEIN.
DR PIR: A02053; GIHUNI.
DR HSSP; P01772; 2IG2.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_V.
DR SMART; SM00406; IgV; 1.
DR Immunoglobulin V region.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 22 96
FT NON_TER 119 119
SQ SEQUENCE 119 AA; 13242 MW; C96935A6E5E165B C8B64;

```

```

alignment_scores:
  Quality: 351.50      Length: 128
  Ratio: 3.515        Gaps: 3
Percent Similarity: 78.125 Percent Identity: 58.594

```

```

alignment_block:
US-09-742-690-1 x HV3I_HUMAN

```

```

Align seg 1/1 to: HV3I_HUMAN from: 1 to: 119

```

```

12  CAGGTGTCAGCTCAGGAGTCAGGAGGAGTGGTCAGGCTGGGGGCTC 61
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
1  GlnValGlnLeuValGlnSerGlyGlyValValGlnProGlyArgSer 17

```

```

62  TCTGAGACTCTCTGTGTCAGCTCGGGACGGCCACAGTGGTCATGTC 111
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
17  rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSer.....A 31
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
112  ACTATGTTATGGGCTGCTCCGCCAGGTTCAGGAAGACGCGTGGT 161
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
31  rGlyThrIleHisTyrValArgGlnAlaProGlyLysGlyLeuGluTyr 47
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
162  GTCGACGCTATTAGTGGAGTGGTAAGACAGACATGTTAAAGACTCC 211
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
48  ValAlaValMetSerTyr***Gly*****LysHisTyrAlaAspSer 64
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
212  GAAGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTT 261
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
64  LAsnGlyArgPheThrIleSerArgAsnAspSerLysAsnThrLeuTyr 81
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
262  TGCAATGACAGCCTGAACCTGAAGATACGGCGGTTTATTATTGTC 311
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
81  euAsnMetAsnSerLeuArgProGluAspThrAlaValTyrTyrCysAla 97
312  GCTGACCGGTCGCCGTGGATGATATTCCTCGCGGTGGTGGTTGACTA 361
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
98  .....ArgIleArgAspThrAlaMet.....PhePheAlaHis 108
362  CTGGGGCCAGGGGACCCAGGTCACCGTCTCCTCA 395
    |||:|||||:|||||:|||||:|||||:|||||:|||||:
108  strpGlyGlnGlyThrLeuValThrValSerSer 119
seq_name: SwissProt_40:HV3T_HUMAN
seq_documentation_block:
ID HV3T_HUMAN STANDARD; PRT; 116 AA.
AC P01781;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region GAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
[1]
SEQUENCE.
RX MEDLINE=75059123; PubMed=4803843;
RA Watanabe S., Barnikol H.U., Horn J., Bertram J., Hilschmann N.;
RT "The primary structure of a monoclonal IgM-immunoglobulin
RT (macroglobulin Gal.), II: the amino acid sequence of the H-chain (mu-
RT type), subgroup H III. Architecture of the complete IgM-molecule."
RL Hoppe-Seyler's Z. Physiol. Chem. 354:1505-1509(1973).
RN [2]
RP REVISION TO THE COMPOSITION OF 28-33.
RX Hilschmann N.;
RL Submitted (JUN-1975) to the PIR data bank.
CC -1- MISCELLANEOUS: THIS MU CHAIN WAS ISOLATED FROM A WALDENSTROM'S
CC MACROGLOBULIN.
DR PIR: A02064; M3HUGL.
DR HSSP; P01772; 2IG2.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IgV; 1.
DR Immunoglobulin V region.
FT NON_TER 116 116
SQ SEQUENCE 116 AA; 12730 MW; 2C67CA9AAAAA1282 CRC64;

```

```

alignment_scores:
  Quality: 348.00      Length: 128
  Ratio: 3.824        Gaps: 2
Percent Similarity: 71.094 Percent Identity: 60.156
alignment_block:

```

```
US-09-742-690-1 x HV3T_HUMAN
Align seg 1/1 to: HV3T_HUMAN from: 1 to: 116
12 CAGGTGCAGCTGCAGGAGTCCAGGGGAGGATGGTGCAGGCTGGGGGCTC 61
   ::::::::::::::::::::|
1 GluValGlnLeuValGluSerGlyGlyAspLeuValGlnProGlyArgSe 17
62 TCTGAGACTCTCCTGTGTCAGGCTCGGGAGCGCCACCATGGTGCATGGTC 111
   ||::::::::::::::::::|
17 rLeuArgLeuSerCysAlaAlaSerGly.....Phe***Phe*** 31
112 ACTATGGTATGGCTGGCTTCCGCCAGGTTCAGGGAAGGAGCGTGAGTTT 161
   ::::::::::::::::::::|
31 **LeuGlyMetThrTrpValArgGlnAlaProGlyGlyLeuGluTrp 47
162 GTCGACGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
   ||::::::::::::::::::|
48 ValAlaAsnLeuLys****GlySer*****TyrValAspSerVa 64
212 GAAGGCGCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
   ||::::::::::::::::::|
64 llysGlyArgPheThrIleSerArgAspAsnAlaLysAsnSerLeuTyrL 81
262 TCAAAATGAACAGCGCTGAACCTGAAGATACGGCGGTTTATTATTGTGCC 311
   ||::::::::::::::::::|
81 euGlnMetAsnSerLeuArgValGluAspThrAlaLeuTyrTyrCysAla 97
312 GCTCGACCGCTCGCGTGGATGATATTTCCCTGCCGTTGGTTTGACTA 361
   ||::::::::::::::::::|
98 ArgGly.....TrpGlyGlyAspTyr 105
362 CTGGGCGCCAGGAGCCAGGTCACCGCTCTCTCA 395
   ||::::::::::::::::::|
105 rTrpGlyGlnGlyThrLeuValThrValSerThr 116
seq_name: SwissProt_40:HV3U_HUMAN
seq_documentation_block:
ID HV3U_HUMAN STANDARD; PRT; 120 AA.
AC P01782;
DT 21-JUL-1986 (Rel. 01, Created)
DE 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region DOB.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=80020921; PubMed=114209;
RA Steiner L.A., Garcia Fardo A., Margolies M.N.;
RT "Amino acid sequence of the heavy-chain variable region of the
RT crystallizable human myeloma protein Dob.";
RL Biochemistry 18:4068-4080(1979).
RN [2]
RP CRYSTALLIZATION.
RX MEDLINE=80020920; PubMed=114208;
RA Steiner L.A., Lopes A.D.;
RT "The crystallizable human myeloma protein Dob has a hinge-region
RT deletion.";
RL Biochemistry 18:4054-4067(1979).
CC -1- MISCELLANEOUS: THIS GAMMA-1 MYELOMA PROTEIN HAS A DELETION IN THE
CC HINGE REGION. THERE ARE NO LIGHT-HEAVY OR INTER-HEAVY CHAIN
CC DISULFIDE BONDS.
CC PIR; A02065; GIHUBD.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
DR Immunoglobulin V region.
KW NON_TER 120
F"
```

```
SQ SEQUENCE 120 AA; 13440 MW; 880DDE307C4B2627 CRC64;
alignment_scores:
  Quality: 347.00 Length: 128
  Ratio: 3.731 Gaps: 2
  Percent Similarity: 72.656 Percent Identity: 58.594
alignment_block:
US-09-742-690-1 x HV3U_HUMAN ..
Align seg 1/1 to: HV3U_HUMAN from: 1 to: 120
12 CAGGTGCAGCTGCAGGAGTCCAGGGGAGGATGGTGCAGGCTGGGGGCTC 61
   ::::::::::::::::::::|
1 GluValGlnLeuValGluSerGlyGlyAspLeuValGlnProGlyArgSe 17
62 TCTGAGACTCTCCTGTGTCAGGCTCGGGAGCGCCACCATGGTGCATGGTC 111
   ||::::::::::::::::::|
17 rLeuArgLeuSerCysAlaAlaSerGly.....PheAsnPheHisG 31
112 ACTATGGTATGGCTGGCTTCCGCCAGGTTCAGGGAAGGAGCGTGAGTTT 161
   ::::::::::::::::::::|
31 lutyAsnMetHisTrpLeuArgGlnGlyProGlyLysGlyProGluTrp 47
162 GTCGACGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGT 211
   ||::::::::::::::::::|
48 ValSerThrIleThrTrpAsnGlyGlySerValLeuTyrAlaAspSerVa 64
212 GAAGGCGCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
   ||::::::::::::::::::|
64 llysGlyArgPheAlaIleSerArgAspAsnAlaGlnLysThrLeuTyrL 81
262 TCAAAATGAACAGCGCTGAACCTGAAGATACGGCGGTTTATTATTGTGCC 311
   ||::::::::::::::::::|
81 euGlnLeuAsnLeuLeuArgProGluAspThrAlaPheTyrTyrCysAla 97
312 GCTCGACCGCTCGCGTGGATGATATTTCCCTGCCGTTGGTTTGACTA 361
   ||::::::::::::::::::|
98 LysGlyTyrIleTrpAsnGlyAsn.....TrpPheAspSe 109
362 CTGGGCGCCAGGAGCCAGGTCACCGCTCTCTCA 395
   ||::::::::::::::::::|
109 rTrpGlyGlnGlyThrLeuValThrValSerSer 120
seq_name: SwissProt_40:HV3C_HUMAN
seq_documentation_block:
ID HV3C_HUMAN STANDARD; PRT; 117 AA.
AC P01764;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region VH26 precursor.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81101090; PubMed=6450418;
RA Matthyssens G., Rabbitts T.H.;
RT "Structure and multiplicity of genes for the human immunoglobulin
RT heavy chain variable region.";
RL Proc. Natl. Acad. Sci. U.S.A. 77:6561-6565(1980).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC ----- J00236; AAA53516.1; -
DR EMBL; M35415; AAA58735.1; -
DR PIR; A02047; H3HU26
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V-III REGION VH26.
FT NON_TER 117
SQ SEQUENCE 117 AA; 12582 MW; E826733F1A3CB0F1 CRC64;

alignment_scores:
    Quality: 346.50      Length: 100
    Ratio: 4.076         Gaps: 1
Percent Similarity: 85.000 Percent Identity: 71.000

alignment_block:
US-09-742-690-1 x HV3K_HUMAN ..

Align seg 1/1 to: HV3K_HUMAN from: 1 to: 117

12 CAGGTGCAGCTGCAGGAGTGTGTCAGGCTGGGGGCTC 61
   ::::|||||  |||||||  |||||||  |||||||  |||||||
20 GluValGlnLeuLeuGluSerGlyGlyLeuValGlnProGlyGlyse 36
   |||||||  |||||||  |||||||  |||||||  |||||||
62 TCTGAGACTCTCTGTGTCAGCTCGGGACGCCACCATGTCGTC 111
   |||||||  |||||||  |||||||  |||||||  |||||||
36 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 50

112 ACTATGTTGGTGGTGGTTCGCCAGGTTCCAGGAGGAGCGTGT 161
   ||:::||||:||||  |||||||  |||||||  |||||||  |||:::
51 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
   |||||||  |||||||  |||||||  |||||||  |||||||
162 GTCCGAGCTATTAGTGTGAGTGTAAAGACATGGTATAAGACTCCGT 211
   ||:::||||:||||  |||||||  |||||||  |||||||  |||||||
67 ValSerAlaIleSerGlySerGlyGlySerThrTrpTrpGlyAspSerVa 83

212 GAAGGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGTTTATC 261
   |||||||  |||||||  |||||||  |||||||  |||||||
83 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 100

262 TGCRAATGAACAGCTGAACCTGAAGATACGGCGTTTATTATGTCGC 311
   |||||||  |||||||  |||||||  |||||||  |||||||
100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 116

seq_name: SwissProt_40:HV3K_HUMAN

seq_documentation_block:
ID HV3K_HUMAN STANDARD; PRT; 126 AA.
AC P01772;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region KOL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=83289131; PubMed=6884994;
RA Schmidt W.E., Jung H.-D., Palm W., Hilschmann N.;
RT "three-dimensional structure determination of antibodies. Primary
RT structure of crystallized monoclonal immunoglobulin IgG1 KOL, I.";
RL Hoppe-Seyler's Z. Physiol. Chem. 364:713-747(1983).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=81072295; PubMed=7441755;
RA Marquart M., Deisenhofer J., Huber R., Palm W.;

```

```

RT *Crystallographic refinement and atomic models of the intact
RT immunoglobulin molecule ".
RT and 1.0-A resolution.".
RL J. Mol. Biol. 141:369-391(1980).
DR PIR; A02055; GIHUKL.
DR PDB; 2FB4; 12-JUL-89.
DR PDB; 2IG2; 12-JUL-89.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; 3D-structure.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 22 96
FT DISULFID 105 110
FT STRAND 3 7
FT STRAND 11 12
FT TURN 14 15
FT STRAND 18 25
FT HELIX 29 31
FT STRAND 34 39
FT TURN 41 42
FT STRAND 46 51
FT TURN 53 54
FT STRAND 58 60
FT TURN 62 67
FT STRAND 68 73
FT TURN 74 77
FT STRAND 78 83
FT HELIX 88 90
FT STRAND 92 99
FT STRAND 106 106
FT TURN 107 108
FT STRAND 109 109
FT STRAND 113 113
FT STRAND 120 124
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 13718 MW; E4D71B52B16F8776 CRC64;

alignment_scores:
    Quality: 346.00      Length: 137
    Ratio: 3.604         Gaps: 3
Percent Similarity: 70.073 Percent Identity: 56.204

alignment_block:
US-09-742-690-1 x HV3K_HUMAN ..

Align seg 1/1 to: HV3K_HUMAN from: 1 to: 126

12 CAGGTGCAGCTGCAGGAGTGTGTCAGGCTGGGGGCTC 61
   |||||||  |||||||  |||||||  |||||||  |||  ||
1 GluValGlnLeuValGluSerGlyGlyValValGlnProGlyArgse 17

62 TCTGAGACTCTCTGTGTCAGCTCGGGACGCCACCATGTCGTCATGTC 111
   |||||||  |||||||  |||||||  |||||||  |||:::
17 rLeuArgLeuSerCysSerSerSerGlyPheIlePheSerSer..... 31

112 ACTATGTTGGTGGTTCGCCAGGTTCCAGGAGGAGCGTGTGTTT 161
   ||:::||||:||||  |||||||  |||||||  |||||||  |||:::
32 ..TyrAlaMetTrpTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47

162 GTCCGAGCTATTAGTGTGAGTGTAAAGACATGGTATAAGACTCCGT 211
   |||||||  |||  :::::  |||  |||||||
48 ValAlaIleIleTrpAspAspGlySerAspGlnHisTyrAlaAspSerVa 64

212 GAAGGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGTTTATC 261
   |||||||  |||||||  |||||||  |||||||  |||:::
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuPheL 81

262 TGCRAATGAACAGCTGAACCTGAAGATACGGCGTTTATTATGTCGC 311
   |||||||  |||||||  |||||||  |||||||  |||||||
81 euGlnMetAspSerLeuArgProGluAspThrGlyValTyrPheCysAla 97

```

262 TGCAAATGACAGCCTCAAAACCCTGAAGTATCGCCGGTTATTATTGTGCC 311
 |||||
 81 euGIinMetAsnSerLeuarGalglaspThrAlavalrrTrcysAla 97
 |||||
 312 GCTCACCGCGTCGCCGTGGAGTAGATAATTCCTGCCGGTGCGTTGACTCPA 361
 |||||
 98ProTgPInPheGlutry 103

 362 CTGGGGCCAGGGGACCCAGGTACCCGGTCTCTCCA 395
 |||||
 103 rTrpGVIGlnGLythrLeuValThrValSerSer 114
 |||||

seq name: SwisssProt 40:HV3F HUMAN

```

seq_documentation_block:
ID HV3F_HUMAN STANDARD; PRT; 115 AA.
AC P01767;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region BUR.
DE Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homipidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE.
RX MEDLINE=78137069; PubMed=416441;
RA Torano A., Putnam F.W.;
RT "Complete amino acid sequence of the alpha 2 heavy chain of a human
RT IgA2 immunoglobulin of the A2m (2) allotype."
RL Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).
CC -1- MISCELLANEOUS: THE SEQUENCE OF THE ALPHA-2, A2M(2) ALLOTYPIC, C
CC REGION OF THIS MELOMA PROTEIN IS ALSO GIVEN.
PR; A02050; A2HUBU.
DR HSP; P01789; IMCP.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam; PF00047; ig_1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin V region.
FT NON_TER 115 115
SO SPOUNCE 115 AA, 12379 MW; 208876A7DF52DCF4 CRC64;

```

alignment_scores:

Quality:	343.50	Length:	128
Ratio:	3.578	Gaps:	4
Percent Similarity:	75.000	Percent Identity:	58.594

Percent Similar

RECOMMENDATION: 1000

```

alignment_block:
US-09-742-690-1 x HV3F HUMAN ..

```

```

En seg 1/1 to: HV3F_HUMAN from: 1 to: 115
12 CAGGTGCAGCTGCAGGAGCTCAGGGGGAGGATTGTGCAGCTGGGGGCTC 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 GluValGlnLeuValGluThrGlyGlyGlyLeuIleGlnProGlyGlySer 17
62 TCTGAGACTCTCTGTGTCAGGCTCGGGACGCCACCAGTCAGTGTCTCATGGTC 111
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrValSer***His... 32
112 ACTATGATGTGGTGGTGTCCGCCAGGTTCACGGGAAGGAGCGTGCAGTTT 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
33 .....SerMetSerTrpValArgGlnAlaProGlyLysAlaLeu***Trp 47
162 GTCCGACCTATTAGTGGAGTGGTAAGAGACATCGTGTAAAGACTCCGT 211
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
48 valSerAlaIle...TyrArgGlyGlyThrTyrTyrAlaAspSerVa 63
212 GAAGGGCCGATTACCACTCTCCAGAGATAACGCCCAAGACTACGGTTTATC 261

```



```
|||||
63 llyGlyArgPheThrIleSerArgAspSerArg***ThrValTyrL 80
|||||
262 TCCAAATGACACCGCTGAACACTGACATACCGCGCTTTATTATTGTGCC 311
|||||
80 euGlnMet***SerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 96
|||||
312 GCTCGACCGCTCGCGTGATGATATTTCCTCGCGGTGGGTGGTCACTA 361
|||||
97 Arg.....AspLeuAlaAlaAlaArgLeuPhe..... 105
|||||
362 CTGGGCGCAGGAGCCAGCTCACCGTCTCCCTCA 395
|||||
106 ....GlyLysGlyThrValThrValSerSer 115
|||||
```

seq_name: SwissProt_40:HV3E_HUMAN

seq_documentation_block:

```
ID HV3E_HUMAN STANDARD; PRT; 120 AA.
AC P01766;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region BRO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77117674; PubMed=65324;
RA Capra J.D., Hopper J.E.;
RT "Comparative studies on monotypic IgM lambda and IgG kappa from an
RT individual patient. III. The complete amino acid sequence of the VH
RT region of the IgM paraprotein.";
RL Immunohistochemistry 13:995-999(1976).
CC -I- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM IGM ISOLATED FROM THE
CC SERUM OF A PATIENT WITH MALIGNANT LYMPHOMA OF THE WALDENSTROM
CC TYPE.
DR PIR: A02049; M3HUB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig: 1.
DR SMART: SM00406; IgV: 1.
DR Immunoglobulin V region.
KW NON_TER 120
FT SEQUENCE 120 AA; 13227 MW; D3F0428F7C2E6410 CRC64;
```

alignment_scores: Quality: 343.50 Length: 128
Ratio: 3.616 Gaps: 4
Percent Similarity: 74.219 Percent Identity: 60.156

alignment_block:

```
US-09-742-690-1 x HV3E_HUMAN ..
Align seg 1/1 to: HV3E_HUMAN from: 1 to: 120
12 CAGGTGCAGCTCAGAGTCAGGGGAGGATGGTCAGGCTGGGGGCTC 61
|||||
1 GluValGlnLeuValGlnSerGlyGlyLeuValGlnProGlyGly 17
|||||
62 TCTGAGACTCTCTGTCAGCTCGGAGCTCGGAGCGCCAGCTGTCATGTC 111
|||||
17 rleuargLeuSerCysAlaAlaSerGlyPheThrPheSer.....T 31
|||||
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGAAGAGCGTGATTT 161
|||||
31 yTyrAsnMetasnrtpValArgGlnValThrGlyLysGlyLeuGluTrp 47
|||||
162 GTTCAGCTATTAGTGGTGGTAAAGACATGGTATAAAGACTCCGT 211
|||||
```

```
48 ValSerAlaIleGlyThrAlaGly...AspGlnTyrTyrAlaAspSerVa 63
212 GAAGGCCGATTCACCATCTCCAGAGATAACGCCAAGACTACCGTTATC 261
|||||
63 llyGlyArgPheThrIleSerArgAsnAspSerLysAsnThrLeuTyrL 80
|||||
262 TCCAAATGACACCGCTGAACACTGACATACCGCGCTTTATTATTGTGCC 311
|||||
80 euAsnMetasnrSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 96
|||||
312 GCTCGACCGCTCGCGTGATGATATTTCCTCGCGGTGGGTGGTCACTA 361
|||||
97 ArgSerProVal.....SerLeuValAspGlyTrpLeuTy 108
|||||
362 C.....TGGGCCAGGGGACC 377
|||||
108 rTyrTyrGlySerValTrpGlyGlnGlyThr 119
|||||
seq_name: SwissProt_40:HV3A_HUMAN
seq_documentation_block:
ID HV3A_HUMAN STANDARD; PRT; 122 AA.
AC P01762;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region TRO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (MVELOMA PROTEIN TRO).
RX MEDLINE=76023781; PubMed=809331;
RA Kratzin H., Altevogt P., Ruban E., Kortt-A., Staroscik K.,
RA Hilschmann N.;
RT "The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro-);
RT II. The amino acid sequence of the H-chain, alpha-type, subgroup III;
RT structure of the complete IgA-molecule.";
RL Hoppe-Seyler's Z. Physiol. Chem. 356:1337-1342(1975).
CC -I- MISCELLANEOUS: THE SEQUENCE OF THE C REGION IS ALSO GIVEN.
DR PIR: A02045; AlHUTR.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig: 1.
DR SMART: SM00406; IgV: 1.
DR Immunoglobulin V region.
KW NON_TER 122
FT SEQUENCE 122 AA; 13472 MW; 2E21A11DA04D80F9 CRC64;
PYRROLIDONE CARBOXYLIC ACID.
```

alignment_scores: Quality: 341.00 Length: 129
Ratio: 3.410 Gaps: 3
Percent Similarity: 77.519 Percent Identity: 52.713

alignment_block:

```
US-09-742-690-1 x HV3A_HUMAN ..
Align seg 1/1 to: HV3A_HUMAN from: 1 to: 122
12 CAGGTGCAGCTCAGAGTCAGGGGAGGATGGTCAGGCTGGGGGCTC 61
|||||
1 GluValGlnLeuValGlnSerGlyGlyGlyLeuValGlyProGlyGly 17
|||||
62 TCTGAGACTCTCTGTCAGCTCGGAGCTCGGAGCGCCAGCTGTCATGTC 111
|||||
17 rleuargLeuSerCysValAlaSerGly.....PheSerPheArg 31
|||||
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGAAGAGCGTGATTT 161
|||||
31 spPheTyrMetSerTrpIleArg***ThrProGlyLysGlyLeu***Trp 47
|||||
```

```

162 GTCGACGCTATTAGTGAGTGGTAAGACAGACATGGTATTAAGACTCCGT 211
||||| ||| ||||| ||||| ||||| |||||
48 ValSerTyrIleGlySerGlySerThrLeuTyrTyrAlaAspSerVa 64
||||| ||| ||||| ||||| ||||| |||||
212 GAAGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTATC 261
||||| ||||| ||||| ||||| ||||| |||||
64 llysglyArgPheThrIleSerArgAspAlaGlnLysSerLeuTyrL 81
||||| ||||| ||||| ||||| ||||| |||||
262 TCCAAATGAACAGCGCTGAACCTGAAGATAGCGCGTTTATTATTGTC 311
||||| ||||| ||||| ||||| ||||| |||||
81 eu***Met***SerLeuArgThr*****ThrAlaValTyrTyrCys 97
||||| ||||| ||||| ||||| ||||| |||||
312 GCTCAGCGGTCGCGGTGGATGATATTCCTCGCGGTT...GGTTTGA 358
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
98 Ala.....Thr*****Phe***TrpSerThrPheSerLeu** 110

359 CTACTGGCGGAGGCCAGCCAGGTCAACCTGCTCTCTCA 395
:||||| ||||| ||||| ||||| ||||| |||||
110 *TyrTrpGly***Gly***LeuValThrValSerSer 122

```

seq_name: SwissProt_40:HV33_MOUSE

```

seq_documentation_block:
ID   HV33_MOUSE          STANDARD;          PRT;          115 AA.
AC   P01802;
DT   21-JUL-1986 (Rel. 01, Created)
DI   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig heavy chain V-III region W3082.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=82099361; PubMed=6798111;
RA   Johnson N., Slankard J., Paul L., Hood L.;
RT   "The complete V domain amino acid sequences of two myeloma inulin-
RL   binding proteins."
RL   J. Immunol. 128:302-307(1982).
CC   -1- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC   BINDS INULIN.
DR   PIR; D92811; AVMS82.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; Ig; 1.
DR   SMART; SM00406; Ig; 1.
KW   Immunoglobulin V region.
FT   DISULFID 22 98
FT   NON_TER 115 115
SQ   SEQUENCE 115 AA; 12887 MW; 9B4517648C121C5A CRC64;

```

alignment_scores:
 Quality: 340.50 Length: 130
 Ratio: 3.510 Gaps: 3
 Percent Similarity: 74.615 Percent Identity: 54.615

alignment_block:
 US-09-742-690-1 x HV33_MOUSE ..

Align seg 1/1 to: HV33_MOUSE from: 1 to: 115

```

12 CAGGTGCACGTCAGGAGTCAGGGGAGGATGGTGCAGCGTGGGGCTC 61
:||||| ||||| ||||| ||||| ||||| |||||
1 GluValLysLeuGlnSerGlyGlyLeuValGlnProGlyGlySe 17

62 TCTGAGACTCTCTGTGACGCTCGGGAGCGCCACCACTGTCATGGTC 111
||||| ||||| ||||| ||||| ||||| |||||
17 rMetLysLeuSerCysValAlaSerGlyPheThrPheSer.....A 31

112 ACTATGGTATGGGTGGTTCGCCAGGTTCACAGGGAAGAGCGTGGT 161

```

```

31 snTyrTrpMetAsnTrpValArgGlnSerProGluLysGlyLeuGluTrp 47
||||| ||||| ||||| ||||| ||||| |||||
162 GTCGACGCTATTAGTGAGTGGTAAGAG.....ACATGGTATAAAGA 205
||||| ||||| ||||| ||||| ||||| |||||
48 ValAlaGluIleArgLeuLysSerHisAsnTyrAlaThrHisTyrAlaG 64
||||| ||||| ||||| ||||| ||||| |||||
206 CTCGCTGAAGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGG 255
||||| ||||| ||||| ||||| ||||| |||||
64 userValLysGlyArgPheThrIleSerArgAspSerLysSerSerV 81
||||| ||||| ||||| ||||| ||||| |||||
256 TTTATCTGCAATGAACAGCGCTGAACCTGAAGATACGCGCGTTTATTAT 305
||||| ||||| ||||| ||||| ||||| |||||
81 alTyrLeuArgMetAsnAsnLeuArgProGluAspThrGlyIleTyrTyr 97
||||| ||||| ||||| ||||| ||||| |||||
306 TGTGCCGCTCGACCGGTCGCGGTGGATGATATTCCTCGCGGTGGGTT 355
||||| ||||| ||||| ||||| ||||| |||||
98 CysThr.....ThrGlyPhe 102

356 TGACTACTGGGCGGAGGCCAGCCAGGTCAACCTGCTCTCTCA 395
| ||||| ||||| ||||| ||||| ||||| |||||
102 eAlaTyrTrpGlyGlnGlyThrLeuValThrValSerAla 115

```

seq_name: SwissProt_40:HV3L_HUMAN

```

seq_documentation_block:
ID   HV3L_HUMAN          STANDARD;          PRT;          119 AA.
AC   P01773;
DT   21-JUL-1986 (Rel. 01, Created)
DI   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   Ig heavy chain V-III region BUR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE (MYELOMA PROTEIN BUR).
RX   MEDLINE=79151016; PubMed=107164;
RA   Putnam F.W., Liu Y.-S.V., Low T.L.K.;
RT   "Primary structure of a human IgA1 immunoglobulin. IV. Streptococcal
RT   IgA1 protease, digestion, Fab and Fc fragments, and the complete
RT   amino acid sequence of the alpha 1 heavy chain."
RT   J. Biol. Chem. 254:2865-2874(1979).
DR   PIR; A02056; A1HUBR.
DR   HSP; P01772; 2IG2.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003596; Ig_v.
DR   Pfam; PF00047; Ig; 1.
DR   SMART; SM00406; Ig; 1.
KW   Immunoglobulin V region; Glycoprotein.
FT   MOD_RES 1 1
FT   DISULFID 22 96
FT   CARBOHYD 28 28
FT   NON_TER 119 119
SQ   SEQUENCE 119 AA; 12981 MW; 12A709A75344D024 CRC64;

```

alignment_scores:
 Quality: 340.50 Length: 127
 Ratio: 3.584 Gaps: 2
 Percent Similarity: 74.803 Percent Identity: 55.906

alignment_block:
 US-09-742-690-1 x HV3L_HUMAN ..

Align seg 1/1 to: HV3L_HUMAN from: 1 to: 119

```

12 CAGGTGCACGTCAGGAGTCAGGGGAGGATGGTGCAGCGTGGGGCTC 61
||||| ||||| ||||| ||||| ||||| |||||
1 GluValGlnLeuValGlnSerGlyGlyValValGlnAlaGlyThrSe 17

62 TCTGAGACTCTCTGTGACGCTCGGGAGCGCCACCACTGTCATGGTC 111

```

```

|||||
17 rLeuArgLeuSerCysThrAlaSerAlaPheAsnLeuSer.....A 31
112 ACTATGGTATGGCGTGGTTCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
|||||
31 sptYrAlaMethIstRpValArgGlnAlaProGlyLysGlyLeu***Trp 47
162 GTCGAGCTATTAGTGGAGTGGTAAAGACATGATGATATAAGACTCCGT 211
|||||
48 ValAlaLeuIleSerTyRcGlySer***ThrTyRAlaAspSerVa 64
212 GAAGGCCGATTCCACATCTCCAGAGATAACGCCAAGACTACGGTTATC 261
|||||
64 lArgGlyArgPheThrIleSerArg***IleSerLys***ThrLeuTyR 81
262 TCAATATGACACCTGGAACCTCAGATACCGCGCTTTATTATTGTCGC 311
|||||
81 eu***MetLysThrLeuArgThrGluAspThrAlaValTyRcysAla 97
312 GCTCAGCCGTCGCGGTGGATGATATTTCCCTGCGCGTGGTGGTTCAC 361
|||||
98 LysLeu.....IleAlaValAlaGlyThrArg***Ph 108
362 CTGGGCCAGGCGCCAGGTCACCGTCTCC 392
|||||
108 etRpGlyGlnGlyThrLeuValThrValSer 118

```

seq_name: SwissProt_40:HV32_MOUSE

seq_documentation_block:

```

ID HV32_MOUSE STANDARD; PRT; 115 AA.
AC P01801;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region J606.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=82099361; PubMed=6798111;
RA Johnson N., Slinkard J., Paul L., Hood L.;
RT "The complete V domain amino acid sequences of two myeloma inulin-
RT binding proteins."
RL J. Immunol. 128:302-307(1982).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM A MYELOMA PROTEIN THAT
CC BINDS INULIN.
DR PIR; C92811; AVMS06.
DR InterPro; IPR003006; Iq_MHC.
DR InterPro; IPR003596; Iq_V.
DR Pfam; PF00047; Iq; 1.
DR SMART; SM00406; Iq; 1.
KW Immunoglobulin V region.
FT DISULFID 22 98 BY SIMILARITY.
FT NON_TER 115 115
SQ SEQUENCE 115 AA; 12810 MW; B67AD6638A121A5F CRC64;

```

alignment_scores:
 Quality: 337.50 Length: 130
 Ratio: 3.479 Gaps: 3
 Percent Similarity: 74.615 Percent Identity: 54.615

alignment_block:

US-09-742-690-1 x HV32_MOUSE ..

Align seg 1/1 to: HV32_MOUSE from: 1 to: 115

```

12 CAGTGTCAGCTCAGGAGTCAGGGGAGGAGGAGTGGTCAGCGTGGGGCTC 61
|||||
1 GluValLysLeuGluSerGlyGlyGlyLeuValGlnProGlyGlyse 17

```

```

62 TCTGAGACTCTCTGTGAGCCTCGGACGCCACAGTGGTGCATGGTC 111
|||||
17 rMetLysLeuSerCysValAlaSerGlyPheThrPheSer.....A 31
112 ACTATGGTATGGCGTGGTTCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
|||||
31 sntYrTrpMetAsnTrpValArgGlnSerProGluLysGlyLeuGluTrp 47
162 GTCGAGCTATTAGTGGAGTGGTAAAGAG.....ACATGGTATATAAGA 205
|||||
48 ValAlaGluIleArgLeuLysSerAsnAsnTyRAlaThrHistyRAlaG 64
206 CTCCGTGAAGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACG 255
|||||
64 uSerValLysGlyArgPheThrIleSerArgAspAspSerLysSerSerV 81
256 TTTATCTGCAATGAACAGCCCTGAACCTGAAGACTACGCCCGTTATTAT 305
|||||
81 altyrLeuGlnMetAsnLeuArgAlaGluAspThrGlyIleTyR 97
306 TGTGCGGCTCGACCGTCCGCGTGGATGATATTTCCCTGCGCGTGGGT 355
|||||
98 CysThr.....ThrglyPh 102
356 TCACCTACTGGGCCAGGACCCAGGTCACCGTCTCTCTCA 395
|||||
102 eaLatyrTrpGlyGlnGlyThrLeuValThrValSerAla 115

```

OM of: US-09-742-690-1 to: PIR_71.* out_format : pfs
Date: Aug 13, 2002 1:43 PM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame-n2p.model -DEV=xlp
-Q/cn2_1/USPTO.spool/US09742690/runat_12082002_094914_102/app_query.fasta_1.623
-DB=PIR_71 -QFMT=fastan -SUFFIX=trp -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -STARF=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09742690_@CGN1_1_73
-NCPU=6 -ICPU=3 -LONGLOC -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-742-690-1
Query length: 558
Database: PIR_71.*
Database sequences: 283138
Database length: 96089334
Search time (sec): 61.480000

score_list:	Strd	Orig	ZScore	EScore	Len	Documentation
PIR2:S70442	+	414.00	629.86	3.9e-27	140	Ig heavy chain precursor V region
PIR2:C36005	+	409.50	624.37	9.2e-27	119	Ig heavy chain V region (30pl)
PIR2:I55673	+	407.50	621.20	1.4e-26	121	Ig heavy chain - human (fragment)
PIR2:S20782	+	407.00	620.24	1.5e-26	124	Ig heavy chain V region - human
PIR2:S50271	+	406.50	617.39	1.7e-26	160	Ig heavy chain precursor - human
PIR2:A60943	+	406.00	617.11	1.8e-26	151	Ig heavy chain precursor V region
PIR2:S31598	+	405.50	617.27	2.0e-26	135	Ig heavy chain V region - human
PIR2:S31686	+	405.00	616.97	2.0e-26	140	Ig heavy chain V region - human
PIR2:S31108	+	404.50	616.79	2.4e-26	119	Ig heavy chain - human
PIR2:S31107	+	403.50	615.28	3.0e-26	119	Ig heavy chain - human
PIR2:S31588	+	402.50	612.42	3.6e-26	140	Ig heavy chain V region - human
PIR2:S31114	+	396.50	604.39	1.7e-25	123	Ig heavy chain - human
PIR2:S31780	+	394.50	599.89	1.7e-25	147	Ig variable region (VDJ) (clone)
PIR2:S48798	+	394.00	600.80	1.9e-25	120	Ig heavy chain V region (anti-s)
PIR2:A49028	+	391.50	596.16	3.1e-25	133	Ig heavy chain V-III region - human
PIR2:S38489	+	390.50	595.03	3.7e-25	127	Ig heavy chain - human (fragment)
PIR2:S31117	+	390.00	594.60	4.1e-25	122	Ig heavy chain - human
PIR2:S31587	+	390.00	593.71	4.1e-25	136	Ig heavy chain V region - human
PIR2:D36005	+	388.50	592.53	5.5e-25	119	Ig heavy chain V region (M43) - human
PIR2:S22657	+	388.50	592.12	5.5e-25	125	Ig V-D-J region (RM) - human (6
PIR2:S22657	+	388.50	591.19	5.5e-25	140	Ig heavy chain precursor V region
PIR2:S31601	+	387.50	590.29	6.7e-25	130	Ig heavy chain V region - human
PIR2:S23624	+	387.50	589.50	6.7e-25	143	Ig heavy chain V region - human
PIR2:G1HUHL	+	386.50	589.36	8.1e-25	121	Ig heavy chain V-III region (H3
PIR2:S31603	+	386.50	588.64	8.1e-25	132	Ig heavy chain V region - human
PIR2:S44111	+	386.00	588.67	8.9e-25	120	Ig heavy chain V-D-J region - human
PIR2:S38864	+	386.00	576.19	9.6e-25	548	Ig epsilon chain C region - mou
PIR2:S31699	+	385.50	587.00	9.9e-25	134	Ig heavy chain V region - human
PIR2:S31120	+	385.00	587.57	1.1e-24	114	Ig heavy chain - human
PIR2:S31116	+	385.00	587.29	1.1e-24	118	Ig heavy chain - human
PIR2:S36278	+	384.50	586.39	1.2e-24	120	Ig heavy chain V region (clone)
PIR2:S31666	+	384.50	585.25	1.2e-24	138	Ig heavy chain V region - human
PIR2:I37781	+	384.50	585.19	1.2e-24	139	Ig variable region (VDJ) (clone)
PIR2:F36005	+	383.50	584.95	1.4e-24	119	Ig heavy chain V region (M49) -
PIR2:S31509	+	383.50	584.67	1.4e-24	123	Ig heavy chain - human
PIR2:S19666	+	382.50	583.29	1.8e-24	121	Ig heavy chain V region (VH3DJH
PIR2:S31112	+	382.00	582.60	1.9e-24	120	Ig heavy chain - human
PIR2:S46391	+	381.00	581.51	2.3e-24	114	Ig heavy chain V region - human
PIR2:S78486	+	380.50	580.54	2.6e-24	117	Ig heavy chain V region (clone)
PIR2:S31105	+	380.00	579.71	2.9e-24	118	Ig heavy chain (subclass IgM) -

PIR2:S55536 + 379.50 578.81 3.1e-24 120 ! Ig heavy chain V region pe20
PIR2:G36005 + 379.00 577.99 3.5e-24 121 ! Ig heavy chain V region (M74
PIR2:S48797 + 379.00 577.52 3.5e-24 128 ! Ig heavy chain V region (ant
PIR2:S55539 + 378.50 577.30 3.8e-24 120 ! Ig heavy chain V region pe24
PIR2:S26794 + 378.50 577.09 3.8e-24 123 ! Ig heavy chain V region - hu

seq_name: pir2:S70442

seq_documentation_block:

Ig heavy chain precursor V region (mu) - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 21-Jan-2000
C:Accession: S70442
Mol. Immunol. 29, 1363-1373, 1992
R:Cuisinier, A.M.; Fumoux, F.; Fougereau, M.; Tonnelle, C.
A:Title: IgM kappa/lambda BBV human B cell clone: an early step of differentiation of
A:Reference number: S70442; MUID:93024508
A:Accession: S70442
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-140 <CUI>
C:Superfamily: immunoglobulin V region: immunoglobulin homology
F:34-117/Domain: immunoglobulin homology <IMM>

alignment_scores:

Quality: 414.00 Length: 129
Ratio: 4.059 Gaps: 2
Percent Similarity: 79.070 Percent Identity: 67.442

alignment_block:

US-09-742-690-1 x S70442 ..
Align seg 1/1 to: S70442 from: 1 to: 140
12 CAGGTGCAGCTCAGAGTCAGGCGGGAGGATGGTGCAGGCTGGGGGCTC 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
20 GlnValGlnLeuValGlnSerGlyGlyValGlnProGlyGlyse 36
62 TCTGAGACTCTCTGTCAGCTCGGCGGCGCCACCGCTGCTGTCGTC 111
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
36 rleuargleuSerCysAlaAlaSerGlyPheThrPheSer.....A 50
112 ACTATGTATGGCTGGTTCCTCCAGGTTCCAGGAGGAGCGTGGT 161
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
50 sntyrGlyMethIstirpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
162 GTCCAGCTATTAGTGGAGTGGTAAAGACATGTTATTAAGACTCCGT 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 ValAlaPheIleArgTyrAspGlySerAsnLysTyrTyrAlaAspSerVa 83
212 GAAGGCGGATTTCACCATCTCCAGAGATAACGCCAAGACTAGCGTTATC 261
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83 llysglyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 100
262 TCGAAATGAACCGCTGAACCTGAAGATPACGCCGCTTTATTATTGTC 311
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 116
312 GTCGACCGCTCGGCGTGATGATATTTCCTCCGCGGTGGGTTCACCTA 361
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
117 Arg.....AspHisIleValGlyAlaThrTyrPheAspTyr 128
362 CTGGGCGGCGGAGGCCAGCTCAGCTCCCTCCCTCAGGA 398
128 rtrpGlyGlnGlyThrLeuValThrValSerSerGly 140

seq_name: pir2:C36005

seq_documentation_block:

Ig heavy chain V region (30pl) - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1990 #sequence_revision 21-Dec-1990 #text_change 16-Aug-1996

C:Accession: C36005
R:Schroeder Jr., H.W.; Wang, J.Y.
Proc. Natl. Acad. Sci. U.S.A. 87, 6146-6150, 1990
A:Title: Preferential utilization of conserved immunoglobulin heavy chain variable gene
A:Reference number: A36005; MUID:90349571
A:Accession: C36005
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-119 <SCH>
C:Cross-references: GB:M18513
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

alignment_scores:
Quality: 409.50 Length: 128
Ratio: 4.054 Gaps: 2
Percent Similarity: 78.906 Percent Identity: 67.188

alignment_block:

US-09-742-690-1 x C36005

Align seg 1/1 to: C36005 from: 1 to: 119

```
12 CAGGTGCAGCTCCAGGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61
   ::::::::::::::::::::
1 GluValGlnLeuLeuGlnSerGlyGlyGlyLeuValGlnProGlyGlySe 17
   ::::::::::::::::::::
62 TCTGAGACTCTCTCTGTCAGCTCGGAGCTCGGGAGCCGCCACGAGTGTGTCATGGTC 111
   ::::::::::::::::::::
17 rleuArgLeuSerCysAlaAlaSerGlyPheThrPheSer..... 31
   ::::::::::::::::::::
112 ACTATGGTATGGCTGGTTCGCCAGGTTCAGGGAAGGAGCGTGAGTTT 161
   ::::::::::::::::::::
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyGlyLeuGluTrp 47
   ::::::::::::::::::::
162 GTCGAGCTATTAGTGGAGTGGTAAAGACATGGTATAAAGACTCCGT 211
   ::::::::::::::::::::
48 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 64
   ::::::::::::::::::::
212 GAAGGCGGATTCACCATCTCCAGACATACCCAGACTAGCTGTTATC 261
   ::::::::::::::::::::
64 llysglyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
   ::::::::::::::::::::
262 TGCAATGAACAGCTGAAACCTGAAGATACGCGCGTTTATTATTGTCGC 311
   ::::::::::::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
   ::::::::::::::::::::
312 GCTCAGCCGGTCCCGTGGATGATATTTCCCTGCCGGTGGGTTTGACTA 361
   ::::::::::::::::::::
98 .....LysAspAlaGlyTrpGlySerGlyPheAspTyr 108
   ::::::::::::::::::::
362 CTGGGCGGAGGAGCCAGGTCACCGTCTCCATCA 395
   ::::::::::::::::::::
108 rTrpGlyGlnGlyThrLeuValThrValSerSer 119
   ::::::::::::::::::::
```

seq_name: p1r2:I55673

seq_documentation_block:

Ig heavy chain - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jan-2000

C:Accession: I55673

R:Knight, G.B.; Agnello, V.; Bonagura, V.; Barnes, J.L.; Panka, D.J.; Zhang, Q.X.

J. Exp. Med. 178, 1903-1911, 1993

A:Title: Human rheumatoid factor cross-idiotypes. IV. Studies on WA Xid-positive IgM with tinct from the 17.109 and G6 Xids.

A:Reference number: I55673; MUID:94065558

C:Accession: I55673

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-121 <RES>

A:Cross-references: GB:M87268; NID:g186197; PIDN:AAC37536.1; PID:g186198
C:Genetics:
A:Gene: GDB:IGHM
A:Cross-references: GDB:120086; OMIM:147020
A:Map position: 14q32.33-14q32.33
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:15-98/Domain: immunoglobulin homology <IMM>

alignment_scores:
Quality: 407.50 Length: 128
Ratio: 3.956 Gaps: 3
Percent Similarity: 80.469 Percent Identity: 67.969

alignment_block:

US-09-742-690-1 x I55673

Align seg 1/1 to: I55673 from: 1 to: 121

```
12 CAGGTGCAGCTCCAGGAGTCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61
   ::::::::::::::::::::
1 GluValGlnLeuLeuGlnSerGlyGlyGlyLeuValGlnProGlyGlySe 17
   ::::::::::::::::::::
62 TCTGAGACTCTCTCTGTCAGCTCGGAGCTCGGCCACGAGTGTGTCATGGTC 111
   ::::::::::::::::::::
17 rleuArgLeuSerCysThrAlaSerGlyPheThrPheSer.....T 31
   ::::::::::::::::::::
112 ACTATGGTATGGCTGGTTCGCCAGGTTCAGGGAAGGAGCGTGAGTTT 161
   ::::::::::::::::::::
31 hrTyrGlyMetSerTrpValArgGlnAlaProGlyGlyGlyLeuGluTrp 47
   ::::::::::::::::::::
162 GTCGAGCTATTAGTGGAGTGGTAAAGACATGGTATAAAGACTCCGT 211
   ::::::::::::::::::::
48 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 64
   ::::::::::::::::::::
212 GAAGGCGGATTCACCATCTCCAGACATACCCAGACTAGCTGTTATC 261
   ::::::::::::::::::::
64 llysglyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
   ::::::::::::::::::::
262 TGCAATGAACAGCTGAAACCTGAAGATACGCGCGTTTATTATTGTCGC 311
   ::::::::::::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
   ::::::::::::::::::::
312 GCTCAGCCGGTCCCGTGGATGATATTTCCCTGCCGGTGGGTTTGACTA 361
   ::::::::::::::::::::
98 AlaAlaProArgHisAlaGly.....SerProPro.....TyrAspTyr 110
   ::::::::::::::::::::
362 CTGGGCGGAGGAGCCAGGTCACCGTCTCCATCA 395
   ::::::::::::::::::::
110 rTrpGlyGlnGlyThrLeuValThrValSerSer 121
   ::::::::::::::::::::
```

seq_name: p1r2:S20782

seq_documentation_block:

Ig heavy chain V region - human

C:Species: Homo sapiens (man)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001

C:Accession: S20782

R:Mortari, F.; Wang, J.; Schroeder, H.W.

submitted to the EMBL Data Library, April 1992

A:Description: Analysis of the IgA and IgG rearranged VH repertoire of human cord blood

A:Reference number: S20782

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-124 <EMBL>

A:Cross-references: EMBL:Z11946; NID:g33897; PIDN:CAA78003.1; PID:g33898

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:15-98/Domain: immunoglobulin homology <IMM>

alignment_scores:

Quality: 407.00 Length: 128
Ratio: 3.913 Gaps: 2
Percent Similarity: 81.250 Percent Identity: 64.844

alignment_block:

US-09-742-690-1 x S20782

Align seg 1/1 to: S20782 from: 1 to: 124

```

12 CAGTGCAGCTGCAGGAGTCAAGGGAGGAGTGTGTCAGGCTGGGGGCTC 61
   ::::::::::::::::::::
1 GluValGlnLeuLeuGluSerGlyGlyLeuValGlnProGlyGlyse 17
   ::::::::::::::::::::
62 TCTGAGACTCTCTGTGTCAGCTCGGGACGCCACCAGCTGTCATGCTC 111
   ::::::::::::::::::::
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSer..... 31
   ::::::::::::::::::::
112 ACTATGTTATGGCTGTTCGCCAGGTTCCAGGAAGGAGCGTGCAGTTT 161
   ::::::::::::::::::::
32 ..TyrAlaMetAsnTrpValArgGlnAlaProGlyGluGlyLeuGluTrp 47
   ::::::::::::::::::::
162 GTCCAGCACTATTAGTGGAGTGGTAAAGACAGATGTTAAAGACTCCGT 211
   ::::::::::::::::::::
48 ValSerThrIleSerGlySerGlyAspSerThrTyrTyrAlaAspSerVa 64
   ::::::::::::::::::::
212 GAAGGGCGGATTCACCACTCCAGAGATAACGCCAAGACTACGGTTTATC 261
   ::::::::::::::::::::
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
   ::::::::::::::::::::
262 TGCATAATCAACAGCTGAACCTGAGATAGATAGCGCGTATTATGTCGC 311
   ::::::::::::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrCysAla 97
   ::::::::::::::::::::
312 GCTGACCGGTCGGCTGATGATATTTCCCTGCGGTTGGGTTTGACTA 361
   ::::::::::::::::::::
98 LysGluArgIleAlaIlePheGlyValValIlePro....HisPheAspTy 113
   ::::::::::::::::::::
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCCTCA 395
   ::::::::::::::::::::
113 rTlpGlyGlnGlyThrLeuValThrValSerSer 124
   ::::::::::::::::::::

```

seq_name: pir2:S05271

seq_documentation_block:

Ig heavy chain precursor - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Aug-1996

C:Accession: S05271; S04602

R:Kishimoto, T.

submitted to the EMBL Data Library, March 1989

A:Reference number: S05270

A:Accession: S05271

A:Molecule type: mRNA

A:Residues: 1-160 <KIS1>

A:Cross-references: EMBL: X14584

R:Kishimoto, T.; Okajima, H.; Okumoto, T.; Taniguchi, M.

Nucleic Acids Res. 17, 4385, 1989

A:Title: Nucleotide sequences of the cDNAs encoding the V-regions of H- and L-chains of

A:Reference number: S04601; MUID: 89296497

A:Accession: S04602

A:Molecule type: mRNA

A:Residues: 1-144 <KIS2>

A:Cross-references: EMBL: X14584

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-160/Product: Ig heavy chain (fragment) #status predicted <MAT>

F:34-117/Domain: immunoglobulin homology <IMM>

alignment_scores:

Quality: 406.50
Ratio: 4.025

Length: 130
Gaps: 1

Percent Similarity: 77.692 Percent Identity: 66.923

alignment_block:

US-09-742-690-1 x S05271

Align seg 1/1 to: S05271 from: 1 to: 160

```

12 CAGTGCAGCTGCAGGAGTCAAGGGAGGAGTGTGTCAGGCTGGGGGCTC 61
   ::::::::::::::::::::
20 GluValGlnLeuLeuGluSerGlyGlyLeuValGlnProGlyGlyse 36
   ::::::::::::::::::::
62 TCTGAGACTCTCTGTGTCAGCTCGGGACGCCACCAGTGTGTCATGCTC 111
   ::::::::::::::::::::
36 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSer.....T 50
   ::::::::::::::::::::
112 ACTATGTTATGGCTGTTCGCCAGGTTCCAGGAAGGAGCGTGCAGTTT 161
   ::::::::::::::::::::
50 hrTyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
   ::::::::::::::::::::
162 CTCGAGCTATTAGTGGAGTGGTAAAGACAGATGTTAAAGACTCCGT 211
   ::::::::::::::::::::
67 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 83
   ::::::::::::::::::::
212 GAAGGGCGGATTCACCACTCCAGAGATAACGCCAAGACTACGGTTTATC 261
   ::::::::::::::::::::
83 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 100
   ::::::::::::::::::::
262 TGCATAATCAACAGCTGAACCTGAGATAGCGCGTATTATGTCGC 311
   ::::::::::::::::::::
100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrCysAla 116
   ::::::::::::::::::::
312 GCTGACCGGTCGGCTGATGATATTTCCCTGCGGTTGGGTTTGACTA 361
   ::::::::::::::::::::
117 LysAlaValAlaArgGlyValIleSerTyrTyrTyrGlyMetAspVa 133
   ::::::::::::::::::::
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCCTCAGGATCT 401
   ::::::::::::::::::::
133 lTrpGlyGlnGlyThrValThrValSerSerGlySer 146
   ::::::::::::::::::::

```

seq_name: pir2:A60943

seq_documentation_block:

Ig heavy chain precursor V region (clone HN.14) - human

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999

C:Accession: A60943; A48165

R:Desai, R.; Spatz, L.; Matsuda, T.; Ilyas, A.A.; Berman, J.E.; Alt, F.W.; Kabat, E.A

J. Neuroimmunol. 30, 245, 1990

A:Title: Molecular cloning of a human immunoglobulin heavy chain variable (V-H) regio

A:Reference number: A60943; MUID: 91036050

A:Accession: A60943

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-151 <DES>

R:Desai, R.; Spatz, L.; Matsuda, T.; Ilyas, A.A.; Berman, J.E.; Alt, F.W.; Kabat, E.A

J. Neuroimmunol. 26, 35-41, 1990

A:Title: Molecular cloning of a human immunoglobulin heavy chain variable (V-H) regio

A:Reference number: A48165; MUID: 90094677

A:Accession: A48165

A:Molecule type: mRNA

A:Residues: 1-36, 'M', 38-62, 'AR', 67-151 <DE2>

C>Note: this sequence has been corrected in reference A60943

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:34-117/Domain: immunoglobulin homology <IMM>

alignment_scores:

Quality: 406.00
Ratio: 3.904

Length: 130
Gaps: 2

Percent Similarity: 80.000 Percent Identity: 65.385

alignment_block:

US-09-742-690-1 x A60943 ..

Align seg 1/1 to: A60943 from: 1 to: 151

```

12 CAGGTGCAGCTCAGGAGTCCAGGGGAGGATGGTGCAGGCTGGGGCTC 61
|||||
20 GlnValGlnLeuValGluSerGlyGlyValValGlnProGlyThrSe 36
|||||
62 TCTGAGACTCTCCTGTGAGCTCGGGACCGGCCACCACTGCTCATGTC 111
|||||
36 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSer.....A 50
|||||
112 ACTATGGTATGGCTGGTCCGCCAGGTTCCAGGAAGAGCGTGAGTTT 161
|||||
50 rGlyGlyMetHisTrpValArgGlnAlaProGlyGlnGlyLeuGluTrp 66
|||||
162 GTCGAGCTATTAGTGGAGTGGTAAAGACAGACATGGTATAAGACTCCGT 211
|||||
67 ValSerValIleSerPheAspGlyGlyThrLysIleTyrAlaAspSerVa 83
|||||
212 GAAGGGCCGATTCCACCTCTCCAGAGATACGCCAAGACTACGGTTTATC 261
|||||
83 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrIleTyrL 100
|||||
262 TCCAAATCAACAGCTGAACCTGAAGATACGGCGGTTTATTATTTGTC 311
|||||
100 euGlnMetAsnSerLeuArgProGluAspThrAlaValTyrTyrCysAla 116
|||||
312 GCTCAGCCGGTCCGGTGGATATATTCCTCGCGGTTGGTTTGACTA 361
|||||
117 LysLysAlaAlaProAlaSerThrGlySer...GlyValAspPheAspTr 132
|||||
362 CTGGGGCCAGGGAGCCAGGTCACCGTCTCCTCAGGATCT 401
|||||
132 rTrpGlyGlnGlyThrLeuValThrValSerSerGlySer 145
|||||

```

seq_name: pir2:S31598

seq_documentation_block:

Ig heavy chain V region - human

C:Species: Homo sapiens (man)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S31598

R:Cuisinier, A.M.; Gauthier, L.; Boublil, L.; Fougereau, M.; Tonnelle, C.

A:Description: Mechanisms that generate human immunoglobulin diversity operate from the

A:Reference number: S31585

A:Accession: S31598

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-135 <CUI>

A:Cross-references: EMBL:Z14170; NID:g31001; PIDN:CAA78539.1; PID:g31002

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:31-114/Domain: immunoglobulin homology <IMM>

alignment_scores:

Quality: 405.50 Length: 129

Ratio: 3.975 Gaps: 3

Percent Similarity: 79.070 Percent Identity: 67.442

alignment_block:

US-09-742-690-1 x S31598 ..

Align seg 1/1 to: S31598 from: 1 to: 135

```

12 CAGGTGCAGCTCAGGAGTCCAGGGGAGGATGGTGCAGGCTGGGGCTC 61
|||||
17 GlnValGlnLeuValGluSerGlyGlyValValGlnProGlyGlySe 33
|||||
62 TCTGAGACTCTCCTGTGAGCTCGGGACCGGCCACCACTGCTCATGTC 111
|||||

```

```

33 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 47
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGAAGAGCGTGAGTTT 161
|||||
48 ..TyrGlyMetHisTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 63
|||||
162 GTCGAGCTATTAGTGGAGTGGTAAAGACAGACATGGTATAAGACTCCGT 211
|||||
64 ValAlaPheIleArgTyrAspGlySerAsnLysTyrTyrAlaAspSerVa 80
|||||
212 GAAGGGCCGATTCCACCTCTCCAGAGATACGCCAAGACTACGGTTTATC 261
|||||
80 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 97
|||||
262 TCCAAATCAACAGCTGAACCTGAAGATACGGCGGTTTATTATTTGTC 311
|||||
97 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 113
|||||
312 GCTCAGCCGGTCCGGTGGATATATTCCTCGCGGTTGGG...TTTGA 358
|||||
114 .....LysIleSerTrpGluValSerArgPheAs 123
|||||
359 CTACTGGGGCCAGGGAGCCAGGTCACCGTCTCTCTCA 395
|||||
123 pTyrTrpGlyGlnGlyThrLeuValThrValSerSer 135
|||||

```

seq_name: pir2:S31686

seq_documentation_block:

Ig heavy chain V region - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999

C:Accession: S31686

R:Cuisinier, A.M.; Gauthier, L.; Boublil, L.; Fougereau, M.; Tonnelle, C.

A:Description: Mechanisms that generate human immunoglobulin diversity operate from

A:Reference number: S31585

A:Accession: S31686

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140 <CUI>

A:Cross-references: EMBL:Z14205; NID:g30969; PIDN:CAA78574.1; PID:g30970

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:34-117/Domain: immunoglobulin homology <IMM>

alignment_scores:

Quality: 405.50 Length: 128

Ratio: 3.975 Gaps: 2

Percent Similarity: 79.688 Percent Identity: 67.188

alignment_block:

US-09-742-690-1 x S31686 ..

Align seg 1/1 to: S31686 from: 1 to: 140

```

12 CAGGTGCAGCTCAGGAGTCCAGGGGAGGATGGTGCAGGCTGGGGCTC 61
|||||
20 GlnValGlnLeuValGluSerGlyGlyGlyLeuValGlnProGlyGlySe 36
|||||
62 TCTGAGACTCTCCTGTGAGCTCGGGACCGGCCACCACTGCTCATGTC 111
|||||
36 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 50
|||||
112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGAAGAGCGTGAGTTT 161
|||||
51 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
|||||
162 GTCGAGCTATTAGTGGAGTGGTAAAGACAGACATGGTATAAGACTCCGT 211
|||||
67 ValSerAlaIleSerGlySerGlySerGlySerThrTyrTyrSerAspSerVa 83
|||||

```


seq_documentation_block:

Ig heavy chain V region - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 23-Jul-1999
 C:Accession: S31588
 R:Cuissinier, A.M.; Gauthier, L.; Boubli, L.; Fougereau, M.; Tonnelie, C.
 submitted to the EMBL Data Library, June 1992
 A:Description: Mechanisms that generate human immunoglobulin diversity operate from the
 A:Reference number: S31585
 A:Accession: S31588
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-140 <CUI>
 A:Cross-references: EMBL:Z14200; NID:g30957; PIDN:CAA78569.1; PID:g30958
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:34-117/Domain: immunoglobulin homology <IMM>

alignment_scores:

Quality: 402.50 Length: 128
 Ratio: 3.946 Gaps: 2
 Percent Similarity: 79.688 Percent Identity: 66.406

alignment_block:

US-09-742-690-1 x S31588 ..

Align seg 1/1 to: S31588 from: 1 to: 140

```

12 CAGGTGACGTCGAGGAGTCAGGGGAGGAGTGGTGCGAGCGTGGGGGTC 61
   ::::::::::::::::::::
20 GluValGlnLeuLeuGluSerGlyGlyGlyLeuValGlnProGlyGlySe 36
   ::::::::::::::::::::
62 TCTGACACTCTCTGTGACGCTCGGCGCGCCACCGTGGTGCATGGTC 111
   ::::::::::::::::::::
36 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 50
   ::::::::::::::::::::
112 ACTATGTTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
   ::::::::::::::::::::
51 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
   ::::::::::::::::::::
162 GTCCGACGCTATTAGTGGAGTGGTAAAGACACATGTTATAAGACATCCGT 211
   ::::::::::::::::::::
67 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 83
   ::::::::::::::::::::
212 GAAGGCGCGATTCCACCATCTCCAGACATACGCCAAGACTACGGTTTATC 261
   ::::::::::::::::::::
83 lLysGlyArgPheThrIleSerArgAspSerLysAsnThrLeuTyrL 100
   ::::::::::::::::::::
262 TGCAAATGAACAGCCCTGAACCTGAAGATACGCCCTTTATTATTGTGC 311
   ::::::::::::::::::::
100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 116
   ::::::::::::::::::::
312 GCTCGACCGCTCGCGTGGATGATATATTCCTCGCGGTGGTGGTTGACTA 361
   ::::::::::::::::::::
117 .....LysAspHisAspTyrSerAsnTyrIleTyrPheAspTyr 129
   ::::::::::::::::::::
362 CTGGGCGCCAGGCGACCCAGTCCACCTCCCTCTCCCTCA 395
   ::::::::::::::::::::
129 rTrpGlyGlnGlyThrLeuValThrValSerSer 140
   ::::::::::::::::::::

```

seq_name: pir2:S31114

seq_documentation_block:

Ig heavy chain - human
 C:Species: Homo sapiens (man)
 C:Date: 02-Dec-1993 #sequence_revision 26-May-1995 #text_change 17-Mar-1999
 C:Accession: S31114
 R:Raaphorst, F.M.; Timmers, E.; Kenter, M.J.H.; van Tol, M.J.D.; Vossen, J.M.; Schuurman
 Eur. J. Immunol. 22, 247-251, 1992
 A:Title: Restricted utilization of germ-line V(H)3 genes and short diverse third comple
 A:Reference number: S31104; MUID:92111633
 A:Accession: S31114

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-123 <RAA>
 A:Cross-references: EMBL:X62963
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotetramer; immunoglobulin
 F:15-98/Domain: immunoglobulin homology <IMM>

alignment_scores:

Quality: 396.50 Length: 128
 Ratio: 3.926 Gaps: 2
 Percent Similarity: 78.906 Percent Identity: 66.406

alignment_block:

US-09-742-690-1 x S31114 ..

Align seg 1/1 to: S31114 from: 1 to: 123

```

12 CAGGTGACGTCGAGGAGTCAGGGGAGGAGTGGTGCGAGCGTGGGGGTC 61
   ::::::::::::::::::::
1 GluValGlnLeuLeuGluSerGlyGlyGlyLeuValGlnProGlyGlySe 17
   ::::::::::::::::::::
62 TCTGACACTCTCTGTGACGCTCGGCGCGCCACCGTGGTGCATGGTC 111
   ::::::::::::::::::::
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
   ::::::::::::::::::::
112 ACTATGTTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161
   ::::::::::::::::::::
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
   ::::::::::::::::::::
162 GTCCGACGCTATTAGTGGAGTGGTAAAGACACATGTTATAAGACATCCGT 211
   ::::::::::::::::::::
48 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 64
   ::::::::::::::::::::
212 GAAGGCGCGATTCCACCATCTCCAGACATACGCCAAGACTACGGTTTATC 261
   ::::::::::::::::::::
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
   ::::::::::::::::::::
262 TGCAAATGAACAGCCCTGAACCTGAAGATACGCCCTTTATTATTGTGC 311
   ::::::::::::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
   ::::::::::::::::::::
312 GCTCGACCGCTCGCGTGGATGATATATTCCTCGCGGTGGTGGTTGACTA 361
   ::::::::::::::::::::
98 LysAlaSerLeuTyrLeuArg.....PheLeuGluTrpLeuPheAspTyr 112
   ::::::::::::::::::::
362 CTGGGCGCCAGGCGACCCAGTCCACCTCCCTCTCCCTCA 395
   ::::::::::::::::::::
112 rTrpGlyGlnGlyThrLeuValThrValSerSer 123
   ::::::::::::::::::::

```

seq_name: pir2:I37780

seq_documentation_block:

Ig variable region (VDJ) (clone T20-11) - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 23-Jul-1999
 C:Accession: I37780; S25474
 R:Demaision, C.; Chastagner, P.; Theze, J.; Zouali, M.
 Proc. Natl. Acad. Sci. U.S.A. 91, 514-518, 1994
 A:Title: Somatic diversification in the heavy chain variable region genes expressed
 A:Reference number: A36876; MUID:94119917
 A:Accession: I37780
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-147 <RES>
 A:Cross-references: EMBL:X67943; NID:g33578; PIDN:CAA48130.1; PID:g33579
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 F:28-111/Domain: immunoglobulin homology <IMM>

alignment_scores:

```

Quality: 394.50      Length: 131
Ratio: 3.757        Gaps: 3
Percent Similarity: 80.153  Percent Identity: 65.649

alignment_block:
US-09-742-690-1 x 137780

Align seg 1/1 to: I37780 from: 1 to: 147

12 CAGGTGCAGCTCAGAGTCAGGGGGAGGATTGGTCAGGCTGGGGGCTC 61
   ::::::::::::::::::::
14 GluValGlnLeuValGlnSerGlyGlyGlyLeuValGlnProGlyGlyse 30
   ::::::::::::::::::::
62 TCTGAGACTCTCTGTGACGCTCGGAGCGCCACCAGTGTGTCATGGTC 111
   ::::::::::::::::::::
30 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 44
   ::::::::::::::::::::
112 ACTATGTATGGCTGGTTCCTCCAGGTTCAGGGAAGAGCGTGAGTTT 161
   ::::::::::::::::::::
45 ..TyrTrpMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 60
   ::::::::::::::::::::
162 GTCGACGATTATTAGTGGAGTGGTAAAGACATGGTATAAGACTCCGT 211
   ::::::::::::::::::::
61 ValAlaAlaSerGlyGlnAspGlySerGluLysTyrTyrAlaAspSerVa 77
   ::::::::::::::::::::
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
   ::::::::::::::::::::
77 LysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 94
   ::::::::::::::::::::
262 TCGAATGAACAGCTGAACCTGAAGATACGGCGCTTTATTATTGTGCC 311
   ::::::::::::::::::::
94 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 110
   ::::::::::::::::::::
312 GCTCGACCGCTCGGCTGGATGATATTCCTGCCGCTTTCACCTA 361
   ::::::::::::::::::::
111 Lys..AspGlyGluGlyTrpGlyLeuTyr..TyrTyrTyrGlyMetAspVa 126
   ::::::::::::::::::::
362 CTGCGGCCAGGGGACCCAGCTCACCCTCTCCCTCAGGATCT 401
   ::::::::::::::::::::
126 lrrpGlyGlnGlyThrThrValThrValSerSerGlySer 139
   ::::::::::::::::::::

seq_name: pir2:S48798

seq_documentation_block:
Ig heavy chain V region (anti-Sm, VH3/Dxp4/JH4b) - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S48798
R:Mahmoudi, M.; Edwards, J.; Cairns, E.; Bell, D.
submitted to the EMBL Data Library, October 1994
A:Description: Molecular characterization of natural human anti-Sm autoantibodies.
A:Reference number: S48797
A:Accession: S48798
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-120 <MAH>
A:Cross-references: EMBL:Z46382; NID:q562324; PIDN:CA86521.1; PID:g1340167
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

alignment_scores:
Quality: 394.00      Length: 129
Ratio: 3.863        Gaps: 3
Percent Similarity: 79.070  Percent Identity: 65.891

alignment_block:
US-09-742-690-1 x S48798

Align seg 1/1 to: S48798 from: 1 to: 120

12 CAGGTGCAGCTCAGAGTCAGGGGGAGGATTGGTCAGGCTGGGGGCTC 61
   ::::::::::::::::::::
14 GluValGlnLeuValGlnSerGlyGlyGlyLeuValGlnProGlyGlyse 30
   ::::::::::::::::::::
62 TCTGAGACTCTCTGTGACGCTCGGAGCGCCACCAGTGTGTCATGGTC 111
   ::::::::::::::::::::
30 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 44
   ::::::::::::::::::::
112 ACTATGTATGGCTGGTTCCTCCAGGTTCAGGGAAGAGCGTGAGTTT 161
   ::::::::::::::::::::
45 ..TyrTrpMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 60
   ::::::::::::::::::::
162 GTCGACGATTATTAGTGGAGTGGTAAAGACATGGTATAAGACTCCGT 211
   ::::::::::::::::::::
61 ValAlaAlaSerGlyGlnAspGlySerGluLysTyrTyrAlaAspSerVa 77
   ::::::::::::::::::::
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
   ::::::::::::::::::::
77 LysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 94
   ::::::::::::::::::::
262 TCGAATGAACAGCTGAACCTGAAGATACGGCGCTTTATTATTGTGCC 311
   ::::::::::::::::::::
94 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 110
   ::::::::::::::::::::
312 GCTCGACCGCTCGGCTGGATGATATTCCTGCCGCTTTCACCTA 361
   ::::::::::::::::::::
111 Lys..AspGlyGluGlyTrpGlyLeuTyr..TyrTyrTyrGlyMetAspVa 126
   ::::::::::::::::::::
362 CTGCGGCCAGGGGACCCAGCTCACCCTCTCCCTCAGGATCT 401
   ::::::::::::::::::::
126 lrrpGlyGlnGlyThrThrValThrValSerSerGlySer 139
   ::::::::::::::::::::

seq_name: pir2:S48798

seq_documentation_block:
Ig heavy chain V region (anti-Sm, VH3/Dxp4/JH4b) - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Jul-1999
C:Accession: S48798
R:Mahmoudi, M.; Edwards, J.; Cairns, E.; Bell, D.
submitted to the EMBL Data Library, October 1994
A:Description: Molecular characterization of natural human anti-Sm autoantibodies.
A:Reference number: S48797
A:Accession: S48798
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-120 <MAH>
A:Cross-references: EMBL:Z46382; NID:q562324; PIDN:CA86521.1; PID:g1340167
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

alignment_scores:
Quality: 391.50      Length: 130
Ratio: 3.764        Gaps: 3
Percent Similarity: 80.000  Percent Identity: 65.385

alignment_block:
US-09-742-690-1 x A49028

Align seg 1/1 to: A49028 from: 1 to: 133

12 CAGGTGCAGCTCAGAGTCAGGGGGAGGATTGGTCAGGCTGGGGGCTC 61
   ::::::::::::::::::::
14 GluValGlnLeuValGlnSerGlyGlyGlyLeuValGlnProGlyGlyse 17
   ::::::::::::::::::::
62 TCTGAGACTCTCTGTGACGCTCGGAGCGCCACCAGTGTGTCATGGTC 111
   ::::::::::::::::::::
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
   ::::::::::::::::::::
112 ACTATGTATGGCTGGTTCCTCCAGGTTCAGGGAAGAGCGTGAGTTT 161
   ::::::::::::::::::::
45 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
   ::::::::::::::::::::
162 GTCGACGATTATTAGTGGAGTGGTAAAGACATGGTATAAGACTCCGT 211
   ::::::::::::::::::::
48 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 64
   ::::::::::::::::::::
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
   ::::::::::::::::::::
64 llysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
   ::::::::::::::::::::
262 TCGAATGAACAGCTGAACCTGAAGATACGGCGCTTTATTATTGTGCC 311
   ::::::::::::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
   ::::::::::::::::::::
312 GCTCGACCGCTCGGCTGGATGATATTCCTGCCGCTTTCACCTA 358
   ::::::::::::::::::::
98 .....LysAspArgGlyPheTrpSerGlyTyrLysAs 108
   ::::::::::::::::::::
359 CTACTGGGGCCAGGGGACCCAGGTTCACCGTCTCCTCA 395
   ::::::::::::::::::::
108 pTyrTrpGlyGlnGlyThrLeuValThrValSerSer 120
   ::::::::::::::::::::

seq_name: pir2:A49028

seq_documentation_block:
Ig heavy chain V-III region - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Jul-1999
C:Accession: A49028
R:Timmers, E.; Kenter, M.; Thompson, A.; Kraakman, M.E.; Berman, J.E.; Alt, F.W.; Sch
Eur. J. Immunol. 21, 2355-2363, 1991
A:Title: Diversity of immunoglobulin heavy chain gene segment rearrangement in B lymph
A:Reference number: A49028; MUID:92008140
A:Accession: A49028
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-133 <TIM>
A:Cross-references: GB:S64471; NID:g236904; PIDN:AAB20011.1; PID:g236905
A:Experimental source: X-linked agammaglobulinemia patients, B lymphoblastoid cell li
A:Note: sequence extracted from NCBI backbone (NCBI:64471, NCBIP:64470)
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>

alignment_scores:
Quality: 391.50      Length: 130
Ratio: 3.764        Gaps: 3
Percent Similarity: 80.000  Percent Identity: 65.385

alignment_block:
US-09-742-690-1 x A49028

Align seg 1/1 to: A49028 from: 1 to: 133

12 CAGGTGCAGCTCAGAGTCAGGGGGAGGATTGGTCAGGCTGGGGGCTC 61
   ::::::::::::::::::::
14 GluValGlnLeuValGlnSerGlyGlyGlyLeuValGlnProGlyGlyse 17
   ::::::::::::::::::::
62 TCTGAGACTCTCTGTGACGCTCGGAGCGCCACCAGTGTGTCATGGTC 111
   ::::::::::::::::::::
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
   ::::::::::::::::::::
112 ACTATGTATGGCTGGTTCCTCCAGGTTCAGGGAAGAGCGTGAGTTT 161
   ::::::::::::::::::::
45 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
   ::::::::::::::::::::
162 GTCGACGATTATTAGTGGAGTGGTAAAGACATGGTATAAGACTCCGT 211
   ::::::::::::::::::::
48 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 64
   ::::::::::::::::::::
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
   ::::::::::::::::::::
64 llysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
   ::::::::::::::::::::
262 TCGAATGAACAGCTGAACCTGAAGATACGGCGCTTTATTATTGTGCC 311
   ::::::::::::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
   ::::::::::::::::::::
312 GCTCGACCGCTCGGCTGGATGATATTCCTGCCGCTTTCACCTA 358
   ::::::::::::::::::::
98 .....LysAspArgGlyPheTrpSerGlyTyrLysAs 108
   ::::::::::::::::::::
359 CTACTGGGGCCAGGGGACCCAGGTTCACCGTCTCCTCA 395
   ::::::::::::::::::::
108 pTyrTrpGlyGlnGlyThrLeuValThrValSerSer 120
   ::::::::::::::::::::

seq_name: pir2:A49028
```

32 . TyrGlyMethIstrpValLeuGlnAlaProGlyLysGlyLeuGlntrp 47
162 GTCACGACTATTAGCTGGAGTGGTAAAGACATAGGTATAAGACTCCGT 211
48 ValAlaValIletrpTyrAspGlySerAsnIlystrTyrAlaAspSerVa 64
212 GAAGGGCGGATTACCATCTCCAGAGATAAGCAAGACTACGGTTTATC 261
64 LlysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
262 TGCAAAATGAACAGCCTGAACCTCAAGATACGGCGTTTATTATGTGCC 311
81 euGlnMetAsnSerLeuArgAlaGluaspThrAlaValTyrIrcys... 96
312 GCTCAGCGGTCCCGGTGGATGATATTCCTCCGGTGGGTGGTTTGACTA 361
97 AlaArgaspArgArgLeu...ThrIleAlaAlaAlaGlyAsnPheAspty 112
362 CTGGGGCCAGGGACCAGGTCCACCGTCTCCTCAGGATCT 401
112 rTrpGlyGlnGlyThrLeuAlaThrValSerSerGlySer 125

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 07:00:02 ; Search time 77.58 Seconds
(without alignments)
1766.737 Million cell updates/sec

Title: US-09-742-690-1

Perfect score: 558

Sequence: 1 tcgagaaaagacaggtgcag.....ctccagtgctgtgtaataag 558

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2.6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2.6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2.6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2.6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	198	35.5	867	4	US-08-862-124-16
2	196.6	35.2	375	4	US-09-240-274-93
3	196.4	35.2	867	4	US-08-862-124-18
4	196	35.1	918	4	US-08-862-124-13
5	196	35.1	918	4	US-08-862-124-15
6	195	34.9	375	4	US-09-240-274-78
7	191	34.2	378	4	US-09-240-274-192
8	185.6	33.3	903	4	US-09-273-839A-12
9	185.6	33.3	913	4	US-09-273-839A-10
10	185.4	33.2	381	4	US-09-240-274-88
11	183	32.8	360	2	US-08-428-197-21
12	183	32.8	360	5	PCT-US93-10555-21
13	182.2	32.7	372	4	US-09-240-274-71
14	182.2	32.7	372	4	US-09-240-274-76
15	182	32.6	378	4	US-09-240-274-79
16	181.4	32.5	360	2	US-08-428-197-19
17	181.4	32.5	360	5	PCT-US93-10555-19
18	180.8	32.4	384	4	US-09-240-274-70
19	180.6	32.4	372	4	US-09-240-274-73
20	180.6	32.4	372	4	US-09-240-274-74
21	180.6	32.4	372	4	US-09-240-274-75
22	180.6	32.4	375	4	US-09-240-274-89
23	180.4	32.3	378	4	US-09-240-274-193
24	179.8	32.2	360	2	US-08-428-197-23
25	179.8	32.2	360	2	US-08-428-197-25
26	179.8	32.2	360	2	US-08-428-197-27
27	179.8	32.2	360	5	PCT-US93-10555-23

Query Match

35.5% ; Score 198 ; DB 4 ; Length 867 ;

ALIGNMENTS

RESULT 1

US-08-862-124-16

; Sequence 16, Application US/08862124

; Patent No. 6207153

; GENERAL INFORMATION:

; APPLICANT: Dan, Michael D.

; APPLICANT: Maiti, Pradip K.

; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS H11, THAT

; TITLE OF INVENTION: SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES ENCODING THE

; TITLE OF INVENTION: FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS AND

; TITLE OF INVENTION: DETECTION OF CANCERS

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morrison & Foerster LLP

; STREET: 755 Page Mill Road

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304-1018

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/862,124

; FILING DATE: 22-MAY-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Lehnhardt, Susan K.

; REGISTRATION NUMBER: 33,943

; REFERENCE/DOCKET NUMBER: 31608-20001.20

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 813-5600

; TELEFAX: (650) 494-0792

; TELEX: 706141

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 867 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; FEATURE:

; NAME/KEY: CDS

; LOCATION: join(1..855, 862..867)

; US-08-862-124-16

28 179.8 32.2 360 5 PCT-US93-10555-25 Sequence 25, Appl
29 179.8 32.2 360 5 PCT-US93-10555-27 Sequence 27, Appl
30 179.4 32.2 406 1 US-08-471-780C-99 Sequence 99, Appl
31 179.4 32.2 406 1 US-08-467-282B-99 Sequence 99, Appl
32 179.4 32.2 406 2 US-08-471-282A-99 Sequence 99, Appl
33 179.4 32.2 406 2 US-08-466-710C-99 Sequence 99, Appl
34 179.4 32.2 406 3 US-08-471-284B-99 Sequence 99, Appl
35 179.4 32.2 406 3 US-08-468-739C-99 Sequence 99, Appl
36 179.2 32.1 384 4 US-09-240-274-185 Sequence 185, App
37 179 32.1 375 4 US-09-240-274-90 Sequence 90, Appl
38 179 32.1 375 4 US-09-240-274-183 Sequence 183, App
39 179 32.1 381 4 US-09-240-274-87 Sequence 87, Appl
40 178.8 32.0 378 4 US-09-240-274-187 Sequence 187, App
41 178.8 32.0 514 3 US-08-545-809A-23 Sequence 23, Appl
42 178.6 32.0 369 2 US-08-428-197-37 Sequence 37, Appl
43 178.6 32.0 369 5 PCT-US93-10555-37 Sequence 37, Appl
44 178 31.9 375 4 US-09-240-274-91 Sequence 91, Appl
45 177.8 31.9 432 1 US-08-026-320A-1 Sequence 1, Appl

; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-862-124-15

Query Match 35.1%; Score 196; DB 4; Length 918;
 Best Local Similarity 70.2%; Pred. No. 8.2e-54;
 Matches 281; Conservative 0; Mismatches 110; Indels 9; Gaps 1;

QY 12 caggtcagctcagagagtcaggggagagattggtcagagctgggggctctcgaactc 71
 Db 441 CAGGTGAGCTGTGGAGTCTGGGGAGGCGGTGTCACAGCTGGGAGGTCCTCGAGACTC 382
 QY 72 tccgtgagctcgggagcgcgcacagtggtcagtggtcactatggtatgggtgcttc 131
 Db 381 TCCTGTGCAGCCTCTGGATTCCTCCCTCAGAG -----CTTGTATGCACCTGGGTC 331
 QY 132 cgcaggttcagggagggagcgtgagttgtcgcagctattaggtgagtgtaagag 191
 Db 330 CGCAGGCTCTAGCAGAGGCGGTGGAGTGGGTGCAGTTATATCATATGATGGAAGCACT 271
 QY 192 acatgtataagactccgtggaagccgattcaccatctccagagataacgcaagact 251
 Db 270 AATACTACGACACCTCCGTGAAGGCGGATTCACCATCTCCAGAGACACTTCCAGAAC 211
 QY 252 acggtttatctgcaaatgaacgcctgaaacctgaagatacgcgcgtttattattgtgc 311
 Db 210 ACGGTGTATCTAAATGAACGCTGAGACTGAGACACGCGTCTCTATTACTGTGG 151
 QY 312 gctcagcggctcgcgtggtgatatttccctgcgggttgggttgaactggggccag 371
 Db 150 AGAGATCAGAGCCTGTGGGTGACTATGACCACTACTACGTTTGGACGCTGGGGCAA 91
 QY 372 gggaccaggtcaccctctcctcagatctcaccatc 411
 Db 90 GGGACCAGGTCACCGTCTCTCAGGATCCGACAAAC 51

RESULT 6

US-09-240-274-78
 ; Sequence 78, Application US/09240274
 ; Patent No. 6255455
 ; GENERAL INFORMATION:
 ; APPLICANT: Siegel, Donald L.
 ; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
 ; FILE REFERENCE: 09596-42U2
 ; CURRENT APPLICATION NUMBER: US/09/240,274
 ; CURRENT FILING DATE: 1999-01-29
 ; EARLIER APPLICATION NUMBER: 60/081,380
 ; EARLIER FILING DATE: 1998-04-10
 ; EARLIER APPLICATION NUMBER: 60/028,550
 ; EARLIER FILING DATE: 1996-10-11
 ; NUMBER OF SEQ ID NOS: 224
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 78
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: anti-Rh(D) chain D03
 US-09-240-274-78

Query Match 34.9%; Score 195; DB 4; Length 375;
 Best Local Similarity 71.5%; Pred. No. 1.1e-53;
 Matches 274; Conservative 0; Mismatches 100; Indels 9; Gaps 1;

QY 13 aggtgcagctcagagtcaggggagagattggtcagagctgggggctctcagactct 72
 Db 2 aggtgcagctcagagtcaggggagagattggtcagagctgggggctctcagactct 61

QY 73 cctgtcagcctcggagcgccacagtggtcatgctactatggtatggctgggttc 132
 Db 62 cctgtcagcctcggagcgccacagtggtcatgctactatggtatggctgggttc 112
 QY 133 gccaggttcagggagggagcgtgagttgtcgcagctattaggtgagtggttaagaga 192
 Db 113 gccaggttcagggagggagcgtgagttgtcgcagctattaggtgagtggttaagaga 172
 QY 193 catggtataaagactccgtggaagccgattcaccatctccagagataacgcaagacta 252
 Db 173 agaatatgcagactccgtggaagccgattcaccatctccagagataacgcaagacta 232
 QY 253 cgtttatctgcaaatgaacgcctgaaacctgaagatacgcgcgtttattattgtgcg 312
 Db 233 cgtgtatctacaaatgaacgcctgagccgagacacgcgtgtgtattactgtcga 292
 QY 313 ctgcagcggctcgcgtggtgatatttccctgcgggttgggttgaactactggggccag 372
 Db 293 gagaagaagtggttcgggagttatcttattggtcggagtttgaactactggggccag 352
 QY 373 ggaaccaggtcaccgtctcctca 395
 Db 353 gaacctggtcaccgtctcctca 375

RESULT 7
 US-09-240-274-192
 ; Sequence 192, Application US/09240274
 ; Patent No. 6255455
 ; GENERAL INFORMATION:
 ; APPLICANT: Siegel, Donald L.
 ; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
 ; FILE REFERENCE: 09596-42U2
 ; CURRENT APPLICATION NUMBER: US/09/240,274
 ; CURRENT FILING DATE: 1999-01-29
 ; EARLIER APPLICATION NUMBER: 60/081,380
 ; EARLIER FILING DATE: 1998-04-10
 ; EARLIER APPLICATION NUMBER: 60/028,550
 ; EARLIER FILING DATE: 1996-10-11
 ; NUMBER OF SEQ ID NOS: 224
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 192
 ; LENGTH: 378
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: anti-Rh(D) antibody clone SH39
 US-09-240-274-192

Query Match 34.2%; Score 191; DB 4; Length 378;
 Best Local Similarity 71.2%; Pred. No. 2.2e-52;
 Matches 270; Conservative 0; Mismatches 100; Indels 9; Gaps 1;

QY 17 gcagctgcagagtcaggggagagattggtcagagctgggggctctcctgagactctctg 76
 Db 9 gctgctcagagcagctcgtggggagcgtgtccagcctggagcctcctcctgagactctctg 68
 QY 77 tgcagcctcgggagcgcgcacacagtggtcatggtcactatggtatgggtgggttcgcca 136
 Db 69 tgcagcgtctggattcaccttcagtag-----ctatggcatgcactgggttcgcca 119
 QY 137 ggttcagggagagagcgtgagtttgcagctattaggtgagtggttaagagacatg 196
 Db 120 ggttcagggagagagcgtgagtttgcagctattaggtgagtggttaagagacatg 179
 QY 197 gtataaagactcgtgaagggcgattcaccatctccagagataacgccaagactcgt 256
 Db 180 atagcagactcgtgaagggcgattcaccatctccagagataacgccaagactcgt 239
 QY 257 ttatctgcaaatgaacagcctgaaacctgaagatacggcgtttattattgtgcccctcg 316

Db 240 gtattacaataaacagacctgagagccgagagacagcgctgtgtattactgtgtcagagaga 299
Qy 317 accggtcccgctggatgatattccctgcgcgtgtgggtttgactactggggccaggggac 376
Db 300 agaagtggttcgggaggtattcttatgtctcgtgaagtttgactactggggccaggggaac 359
Qy 377 ccaggtccaccgtctctctca 395
Db 360 cctggctaccgtctctctca 378

RESULT 8

US-09-273-839A-12
; Sequence 12, Application US/09273839A
; Patent No. 6329156
; GENERAL INFORMATION:
; APPLICANT: Cirino, Nick M
; APPLICANT: Jackson, Paul J
; APPLICANT: Lehnert, Bruce E
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface
; FILE OF INVENTION: Receptors
; FILE REFERENCE: S-89,662
; CURRENT APPLICATION NUMBER: US/09/273,839A
; CURRENT FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 903
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-273-839A-12

Query Match 33.3%; Score 185.6; DB 4; Length 903;
Best Local Similarity 71.8%; Pred. No. 1.9e-50;
Matches 277; Conservative 0; Mismatches 94; Indels 15; Gaps 2;
Qy 12 cagggtcagctcagagtcagggggagattggtcaggctgggggctctctgagactc 71
Db 102 cagggtcagctcagagtcagggggagattggtcaggctgggggctctctgagactc 161
Qy 72 tctgtgcagctcggagcgcgcacagtggtcagtcagtggtcagtcagtggtcagtc 131
Db 162 tctgtgcagctcggagcgcgcacagtggtcagtcagtggtcagtcagtggtcagtc 212
Qy 132 cccaggttcaggagagcgtgattgttcgcagctattagtggtgagtggtgaaagag 191
Db 213 cgcaggttcaggagagcgtgattgttcgcagctattagtggtgagtggtgaaagag 272
Qy 192 acatggtataaagactccgtgaaggccgattccaccatctccagagataacgccaagact 251
Db 273 acatactacagactccgtgaaggccgattccaccatctccagagataacgccaagac 332
Qy 252 acggtttatctcaaatgaacacgcctgaacccctgaagatacagccgctttattgtgcc 311
Db 333 acgctgtattctcaaatgaacacgcctgaagagctgagagacagcgtgtgtattactgtg-- 390
Qy 312 gctcagccgtccgctgagtagatatttccctgcgcgttgggtttgactactggggccag 371
Db 391 ----tgaagatctccagcttgagcgtcacaattgggggtatttactactggggccag 446
Qy 372 gggaccaggtccagctctctcagg 397
Db 447 ggaacctggaccgtctctcagg 472

RESULT 9

US-09-273-839A-10
; Sequence 10, Application US/09273839A
; Patent No. 6329156
; GENERAL INFORMATION:
; APPLICANT: Cirino, Nick M
; APPLICANT: Jackson, Paul J

; APPLICANT: Lehnert, Bruce E
; TITLE OF INVENTION: Disruption of Anthrax Toxin Binding to Cell Surface
; FILE OF INVENTION: Receptors
; FILE REFERENCE: S-89,662
; CURRENT APPLICATION NUMBER: US/09/273,839A
; CURRENT FILING DATE: 1999-03-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 913
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-273-839A-10

Query Match 33.3%; Score 185.6; DB 4; Length 913;
Best Local Similarity 71.8%; Pred. No. 1.9e-50;
Matches 277; Conservative 0; Mismatches 94; Indels 15; Gaps 2;
Qy 12 cagggtcagctcagagtcagggggagattggtcaggctgggggctctctgagactc 71
Db 103 cagggtcagctcagagtcagggggagattggtcaggctgggggctctctgagactc 162
Qy 72 tctgtgcagctcggagcgcgcacagtggtcagtcagtggtcagtcagtggtcagtc 131
Db 163 tctgtgcagctcggagcgcgcacagtggtcagtcagtggtcagtcagtggtcagtc 213
Qy 132 cgcaggttcaggagagcgtgattgttcgcagctattagtggtgagtggtgaaagag 191
Db 214 cgcaggttcaggagagcgtgattgttcgcagctattagtggtgagtggtgaaagag 273
Qy 192 acatggtataaagactccgtgaaggccgattccaccatctccagagataacgccaagact 251
Db 274 acatactacagactccgtgaaggccgattccaccatctccagagataacgccaagac 333
Qy 252 acggtttatctcaaatgaacacgcctgaacccctgaagatacagccgctttattgtgcc 311
Db 334 acgctgtattctcaaatgaacacgcctgaagagctgagagacagcgtgtgtattactgtg-- 391
Qy 312 gctcagccgtccgctgagtagatatttccctgcgcgttgggtttgactactggggccag 371
Db 392 ----tgaagatctccagcttgagcgtcacaattgggggtatttactactggggccag 447
Qy 372 gggaccaggtccagctctctcagg 397
Db 448 ggaacctggaccgtctctcagg 473

RESULT 10

US-09-240-274-88
; Sequence 88, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF
; FILE REFERENCE: 09596-4202
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 88
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain D14
US-09-240-274-88


```
; NAME: Howells, Stacy L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD-2630
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Spa3-08
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..360
; PCT-US93-10555-21

Query Match          32.8%; Score 183; DB 5; Length 360;
Best Local Similarity 72.2%; Pred. No. 8.4e-50;
Matches 275; Conservative 0; Mismatches 85; Indels 21; Gaps 2;

Qy 12 caggtgcagctgcaggaggtcaggggagattggtgcaggctggggctctctgagactc 71
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1  CAGGTGAACCTGCTCGAGTCTGGGGAGGATTGTTACAGCCTGGGGGCTCCCTGAGACTC 60

Qy 72 tctgtgcagctcgggagcgcacacagtcagtgatgctactatggtatgggtgggttc 131
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TCCTGTGAGCCTCTGATTACCTTTACGAGCCA-----TGCCATGAGCTGGGTC 111

Qy 132 cgcaggtctcagaaagacgctgagttgtgcagctattagtgagtggttaagag 191
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 112 GCCAGGCTCAGGAAGGGCTGGAGTGGGTCCTCAGATATAGTGCAGTGGTGAGC 171

Qy 192 acatggtataaagactcgtgaaggcgcagattcacatctccagagataaacgccaagact 251
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 172 ACATATTATGACACTCCGTCGAAGGGCGGTTCACCATCTCCAGAGACAATTCCAAGAAC 231

Qy 252 acggtttatctgaaataaagacgctgaaacctgaagatacagccgctttattattgtgcc 311
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 232 ACCTGTATTGTTGAAATGAACAGCCTGAGAGCCGAGACACGGCCTTATATTACTGTGCG 291

Qy 312 gctgcagcgtcccgctgagatgatttccctgcggttggtgttgactactggggccag 371
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 292 TCCAGCGCGGGTGGGGCTACCTTCCC-----TTGACTACTGGGGCCAG 339

Qy 372 gggaccagggtcacgctctcc 392
      ||||| ||||| ||||| |||||
Db 340 GGAACCTGGTCACCGTCTCC 360

RESULT 13
US-09-240-274-71
; Sequence 71, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: RH(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-4202
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 71
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain C10
US-09-240-274-76
; Sequence 76, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: RH(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-4202
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain C10
US-09-240-274-76

Query Match          32.7%; Score 182.2; DB 4; Length 372;
Best Local Similarity 71.3%; Pred. No. 1.6e-49;
Matches 273; Conservative 0; Mismatches 98; Indels 12; Gaps 2;

Qy 13 aggtgcagctgcaggagtcaggaggagattggtgcaggctggggctctctgagactct 72
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2  aggtgcagctgctcgagctctggggagggcgtggtccagcctgggaggtccctgagactct 61

Qy 73 cctgtgcagctcgggagcgcgccacagtggtcagtggtcactatggtatgggtgggttccc 132
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 cctgtgcagctcgtgattctcttcagtag-----ctatggcagtcacgtgggtccc 112

Qy 133 gccaggttccagggaaggagcgtgagttgtgcagcagctattagtggtggtggttaagaga 192
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 113 gccaggttccagggaaggagcgtgagttgtgcagctattatcatatgatgagacatcata 172

Qy 193 catggtataaagactcctggaaggccgattcacatctccagagataaacccaagacta 252
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 173 aaaaactatgcagactcctggaaggccgattcacatctccagagataaacccaagaaaa 232

Qy 253 cggtttatctcgaatgaacagcctgaacacctgaagatacacgcccgtttattattgtgccg 312
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 233 cgtgtgacctgcgaatgaacagcctgagacctgagacacggtgtatattactgtgcga 292

Qy 313 ctgcagccggtccgctggatgatatttccctgcggttggtgttgactactggggccagg 372
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 293 acctaaagggggaagtaactcgtctgc---gtctgttcccttgatctatctggggccag 349

Qy 373 ggaacccagggtcacgctctccca 395
      ||||| ||||| ||||| |||||
Db 350 ggacaatggtcacgctctctcca 372

RESULT 14
US-09-240-274-76
; Sequence 76, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: RH(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; FILE REFERENCE: 09596-4202
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain C10
US-09-240-274-76

Query Match          32.7%; Score 182.2; DB 4; Length 372;
Best Local Similarity 71.3%; Pred. No. 1.6e-49;
Matches 273; Conservative 0; Mismatches 98; Indels 12; Gaps 2;

Qy 13 aggtgcagctgcaggagtcaggaggagattggtgcaggctggggctctctgagactct 72
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2  aggtgcagctgctcgagctctggggagggcgtggtccagcctgggaggtccctgagactct 61
```

```
Qy 73 cctgtgacgctcggagcgccgacacagtggtcactatggtatgggtggttcc 132
|||||
Db 62 cctgtgacgctcggatctcttcagtag-----ctatggcatgcactgggtcc 112
|||||
Qy 133 gccaggttcagggaagagcgctgagttgtctgcagctattagtgagtggttaagaga 192
|||||
Db 113 gccaggttcagggaagagcgctgagttgtctgcagttatatcatatgatggacatcata 172
|||||
Qy 193 catggtataaaagactccgtgaagggccgattccaccatctccagagataacgccaaagacta 252
|||||
Db 173 aaaactatgcagactccgtgaagggccgattccaccatctccagagataacgccaaagaaa 232
|||||
Qy 253 cggtttatctgcaaatgaacagcctgaacctgaagatacagccgctttattattgtgcg 312
|||||
Db 233 cgcgtacctgcaaatgaacagcctgaacctgaagacagcgtgtatatattactgtgcga 292
|||||
Qy 313 ctgcacgggtccgctggtgatgatatttccctgccggttggtttgactactcggggccagg 372
|||||
Db 293 acctaaagggggaagtaactcgtgtgc---gtctgttccctttgatctcgtggggccag 349
|||||
Qy 373 ggaaccaggtcacccgtctctca 395
|||||
Db 350 ggaattggtcacccgtctctca 372
|||||
```

RESULT 15

```
US-09-240-274-79
; Sequence 79, Application US/09240274
; Patent No. 6255455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF
; FILE REFERENCE: 09596-4202
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 79
; LENGTH: 378
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain D04
US-09-240-274-79
```

```
Query Match 32.6%; Score 182; DB 4; Length 378;
Best Local Similarity 71.0%; Pred. No. 1.8e-49;
Matches 274; Conservative 0; Mismatches 100; Indels 12; Gaps 2;

Qy 13 aggtcagctcagagtcagagcagagagattggtcagcgtggggctctctgagactct 72
|||||
Db 2 aggtcagctcgtcagtcggtggggagcgctgcccagccctcggggaggtccctgagactct 61
|||||
Qy 73 cctgtgacgctcggagcgcaccagtggtcatggtcactatggtatggctggttcc 132
|||||
Db 62 cctgtgagcgtctggttcagctcagag-----ctatggcatgcactgggtcc 112
|||||
Qy 133 gccaggttcagggaagagcgtgagttgtctgcagcgtattagtgagtggttaagaga 192
|||||
Db 113 gccaggttcagggaagagcgtgagttgtctgcagatatacgtttgatggagtaata 172
|||||
Qy 193 catggtataaaagactccgtgaagggccgattccaccatctccagagataacgccaaagacta 252
|||||
Db 173 aaagattatgcagactccgtgaagggccgattccaccatctccagagataacgccaaagacta 232
|||||
Qy 253 cggtttatctgcgaatgaacagcctgaaacctgaagatacagcgcgtttattattgtgc-- 310
|||||
```

```
Db 233 cgttgtatcttcaaatgaacagcctgagagccgagagatacagcgtgtgtattattgtgcga 292
|||||
Qy 311 -cgctcagaccggtccgcgtggaatgatatattccctgccggttgggtttgactactcggggcc 369
|||||
Db 293 gagattgagaggtgcggcccttttagtggtcgtggttaagtgcttttgatatactcggggcc 352
|||||
Qy 370 aggggacccaggtcaccgctctctctca 395
|||||
Db 353 aagggaacaatggtcacccgtctctca 378
|||||
```

Search completed: August 13, 2002, 12:31:44
Job time: 19902 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 09:24:07 ; Search time 273.29 Seconds
(without alignments)
3505.574 Million cell updates/sec

Title: US-09-742-690-1
Perfect score: 558
Sequence: 1 tcgagaagaacaggtgcag.....ctccacgtgtctgtaataag 558

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID55/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID55/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	385.4	69.1	440	22	AA10052
2	384.8	69.0	1154	22	AA10057
3	384.8	69.0	1172	22	AA10058
4	299.2	53.6	384	20	AA59602
5	291.2	52.2	714	20	AA59612
6	291.2	52.2	717	20	AA59614
7	291.2	52.0	750	20	AA59616
8	290.2	52.0	708	20	AA59617
9	232.8	41.7	375	20	AA578160

10	219.8	39.4	378	20	AA578159	Llama anti-carbazo
11	209.6	37.6	354	20	AA59605	PstI/BstEII fragme
12	209.2	37.5	369	20	AA59607	PstI/BstEII fragme
13	208.4	37.3	351	20	AA59604	PstI/BstEII fragme
14	208.4	37.3	351	21	AA000365	Plasmid pUR4601 en
15	208	37.3	462	22	AA10055	Anti-GUS (Clone18)
16	208	37.3	471	22	AA10056	Anti-GUS VH region
17	201.4	36.1	456	11	AA03610	Sequence encoding
18	199.8	35.8	375	11	AA03608	Sequence encoding
19	198	35.5	867	19	AAV10119	Human HII-scfv con
20	198	35.5	867	22	AA04539	Human monoclonal a
21	197.2	35.3	521	19	AAV17546	Clone CA05 gene.
22	196.6	35.2	375	22	AAH68638	Human anti-Rh(D) c
23	196.6	35.2	524	19	AAV17545	Clone CA04 gene.
24	196.4	35.2	867	22	AA04540	Human HII-single c
25	196	35.1	918	19	AAV10118	Human HII-scfv con
26	196	35.1	918	22	AA04537	Human monoclonal a
27	196	35.1	918	22	AA04538	Human HII-single c
28	195.4	35.0	1458	22	AA52625	Human cDNA encodin
29	195	34.9	375	22	AAH68623	Human anti-Rh(D) c
30	195	34.9	1413	21	AAA46898	DNA encoding the h
31	193.4	34.7	672	20	AA59613	Plasmid pUR4619 fr
32	193.4	34.7	672	21	AA000371	Saccharomyces fusi
33	191.4	34.3	675	20	AA59615	Plasmid pUR4621 fr
34	191	34.2	378	22	AAH68694	Human anti-Rh(D) a
35	190.4	34.1	738	21	AA255614	Internalising anti
36	188.6	33.8	372	22	AA575585	Human anti-HER2/ne
37	188.4	33.8	1741	22	AA522531	Human cDNA encodin
38	187.6	33.6	681	22	AAH30055	TRO005 heavy chain
39	187.2	33.5	366	22	AA575589	Human anti-HER2/ne
40	187.2	33.5	1644	22	AA522593	Human cDNA encodin
41	187	33.5	750	22	AA500212	DNA encoding monoc
42	187	33.5	753	22	AA500211	DNA encoding monoc
43	187	33.5	909	14	AA043772	Sequence encoding
44	185.6	33.3	603	22	AA585221	Nucleotide sequenc
45	185.4	33.2	381	22	AAH68633	Human anti-Rh(D) c

ALIGNMENTS

RESULT 1	
AA10052	AA10052 standard; DNA; 440 BP.
ID	AA10052 standard; DNA; 440 BP.
XX	
AC	AA10052;
XX	
DT	18-SEP-2001 (first entry)
XX	
DE	Llama HCV33 antibody heavy chain variable domain with myc tag DNA.
XX	
KW	Camelid; llama; heavy chain immunoglobulin; pathogen resistance; metabolism modulator; passive immunisation; heavy chain variable domain;
KW	VH; anti-RR6 antibody; HCV33; azo-dye; ds.
XX	
OS	Chimeric - Llama guanicoe llama.
OS	Chimeric - Unidentified.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..420
FT	/*tag= a
FT	/product= "Llama HCV33 antibody heavy chain variable domain with myc peptide"
FT	/note= "Does not include start codon"
FT	/partial
XX	
XX	EP1118669-A2.
PD	25-JUL-2001.
XX	
PF	08-DEC-2000; 2000EP-0310997.
XX	
PR	17-DEC-1999; 99EP-0310188.


```
CC azo-dye) antibody (designated as HCV33) heavy chain variable domain (VH)
CC linked to the non-classical heavy chain constant regions denoted as
CC hinge-CH2-CH3.
XX
SQ Sequence 1154 BP; 293 A; 341 C; 321 G; 199 T; 0 other;

Query Match          69.0%; Score 384.8; DB 22; Length 1154;
Best Local Similarity 99.5%; Pred. No. 1.2e-95;
Matches 386; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 aggtgcagctgcaggaactcaggagggagagattggtgcaggctgggggctctcagactct 72
Db 7 aggtgcagctgcaggaactcaggagggagagattggtgcaggctgggggctctcagactct 66
QY 73 cctgtgcagctcgggagcgccagccagtggtcactatggtgctcactatggtatggctggtcc 132
Db 67 cctgtgcagctcgggagcgccagccagtggtcactatggtgctcactatggtatggctggtcc 126
QY 133 gccaggttccagggaaggagcgctgagttgtcgcagctattaggtgagtggttaagaga 192
Db 127 gccaggttccagggaaggagcgctgagttgtcgcagctattaggtgagtggttaagaga 186
QY 193 catggtataagactccgtgaaggccgattccaccatccagagataacgccaaagacta 252
Db 187 catggtataagactccgtgaaggccgattccaccatccagagataacgccaaagacta 246
QY 253 cggttatctgcaaatgaacagcctgaaactgaacgtgaacgtggtttatttattgtgccg 312
Db 247 cggttatctgcaaatgaacagcctgaaactgaacgtggtttatttattgtgccg 306
QY 313 ctgcagcggctcgcgtggtgatatttccctgcgggttggtggttgcactactggggccagg 372
Db 307 ctgcagcggctcgcgtggtgatatttccctgcgggttggtggttgcactactggggccagg 366
QY 373 ggaccaggtcacgcgtctcctcaggatc 400
Db 367 ggaccaggtcacgcgtctcctcagaacc 394

RESULT 3
AAD10058
ID AAD10058 standard; DNA; 1172 BP.
XX
AC AAD10058;
XX
DT 18-SEP-2001 (first entry)
XX
DE Llama HCV33 VH region with ER retention signal and hinge-CH2-CH3 DNA.
XX
KW Camelid; llama; heavy chain immunoglobulin; pathogen resistance;
KW metabolism modulator; passive immunisation; heavy chain variable domain;
KW VH; anti-RR6 antibody; HCV33; azo-dye; ds.
XX
OS Chimeric - Lama guanicoe glama.
OS Chimeric - Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 3..1154
FT /tag= a
FT /product= "Llama HCV33 heavy chain variable domain with
FT heavy chain constant regions and ER retention signal"
FT misc_feature 3..389
FT /tag= b
FT /note= "Corresponds to Llama HCV33 heavy chain
FT variable domain"
FT misc_feature 390..476
FT /tag= c
FT /note= "Corresponds to hinge region DNA"
FT misc_feature 477..806
FT /tag= d
FT /note= "Corresponds to heavy chain constant region
FT CH2 DNA"
```

```
FT misc_feature 807..1133
FT /tag= e
FT /note= "Corresponds to heavy chain constant region
FT CH3 DNA"
FT misc_feature 1134..1151
FT /tag= f
FT /note= "Corresponds to ER retention signal DNA"
XX
EP1118669-A2.
XX 25-JUL-2001.
XX 08-DEC-2000; 2000EP-0310997.
XX 17-DEC-1999; 99EP-0310188.
XX (UNIL ) UNILEVER PLC.
XX (UNIL ) UNILEVER NV.
XX Frenken LGJ, Van Der Logt CPE, Jobling SA, Teh Y;
XX WPI: 2001-427157/46.
XX P-PSDB; AAE05287.
XX
XX Modifying a plant to produce an antibody useful for increasing pathogen
XX resistance or to modulate metabolism comprises introducing a DNA
XX sequence encoding a heavy chain immunoglobulin linked to a peptide that
XX targets a cellular compartment -
XX Example 12; Fig 19; 81pp; English.
XX
XX The present invention relates to a method for modifying a plant to
XX produce an antibody or an active fragment or derivative, or a protein
XX functional equivalent, in a cellular compartment comprising introducing
XX a DNA sequence encoding a heavy chain immunoglobulin, where the DNA is
XX linked to promoters and provided with an additional sequence encoding a
XX peptide capable of targeting heavy chain immunoglobulin to a cellular
XX compartment. The method is used for producing a heavy chain
XX immunoglobulin or an active fragment or derivative, or a protein that is
XX functionally equivalent for increasing the pathogen resistance in a plant
XX or to modulate metabolism in a plant. Under some circumstances it may be
XX desirable to retain the antibody product with the plant rather than
XX extracting and isolating the product. In particular, edible selected
XX antigens may be used in a method of passively immunising an animal,
XX preferably human, against the antigen, e.g., pathogenic organisms. The
XX present DNA sequence encodes a llama (camelid) anti-RR6 (RR6 is an
XX azo-dye) antibody (designated as HCV33) heavy chain variable domain (VH)
XX linked to the non-classical heavy chain constant regions denoted as
XX hinge-CH2-CH3 and endoplasmic reticulum (ER) retention signal.
XX
XX Sequence 1172 BP; 300 A; 343 C; 326 G; 203 T; 0 other;

Query Match          69.0%; Score 384.8; DB 22; Length 1172;
Best Local Similarity 99.5%; Pred. No. 1.2e-95;
Matches 386; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 aggtgcagctgcaggaactcaggagggagagattggtgcaggctgggggctctcagactct 72
Db 7 aggtgcagctgcaggaactcaggagggagagattggtgcaggctgggggctctcagactct 66
QY 73 cctgtgcagctcgggagcgccagccagtggtcactatggtgctcactatggtatggctggtcc 132
Db 67 cctgtgcagctcgggagcgccagccagtggtcactatggtgctcactatggtatggctggtcc 126
QY 133 gccaggttccagggaaggagcgctgagttgtcgcagctattaggtgagtggttaagaga 192
Db 127 gccaggttccagggaaggagcgctgagttgtcgcagctattaggtgagtggttaagaga 186
QY 193 catggtataagactccgtgaaggccgattccaccatccagagataacgccaaagacta 252
Db 187 catggtataagactccgtgaaggccgattccaccatccagagataacgccaaagacta 246
```

Qy 253 cggtttatctgcaaatgaacagcctgaacctgaagatacggcgggtttattattgtgccc 312
 |||||
 Db 247 cggtttatctgcaaatgaacagcctgaacctgaagatacggcgggtttattattgtgccc 306
 |||||
 Qy 313 ctgacacggctccgctggtgatatttccctgcccgggttggtttgactactcggggccagg 372
 |||||
 Db 307 ctgacacggctccgctggtgatatttccctgcccgggttggtttgactactcggggccagg 366
 |||||
 Qy 373 ggaccaggctcacgctctccaggatc 400
 |||||
 Db 367 ggaccaggctcacgctctccaggatc 394
 |||||

RESULT 4
 AAX59602
 ID AAX59602 standard; DNA; 384 BP.
 XX
 AC AAX59602;
 XX
 DT 21-JUL-1999 (first entry)
 XX
 DE PstI/BstEII fragment of plasmid pUR4538.
 XX
 KW Multivalent antigen binding protein; single polypeptide chain;
 KW single domain binding site; diagnosis; therapy; targeting;
 KW immunoassay; cross-linking; agglutination; purification;
 KW phage inactivation; detergent; ss.
 XX
 OS Synthetic.
 XX
 FN WO9923221-A2.
 XX
 PD 14-MAY-1999.
 XX
 PF 27-OCT-1998; 98WO-EP06991.
 XX
 PR 27-OCT-1997; 97EP-0308538.
 XX
 PA (UNIL) UNILEVER NV.
 PA (UNIL) UNILEVER PLC.
 XX
 PI Frenken LGJ, Howell S, Ledebøer AM, Van Der Logt CPE;
 XX
 DR WPI; 1999-313342/26.
 DR P-PSDB; AAY15398.
 XX
 PT Multivalent antigen binding protein useful for inactivation of
 PT (bacteriophages or viruses
 XX
 PS Disclosure; Fig 2; 89pp; English.
 XX
 CC The specification describes multivalent antigen binding proteins
 CC comprising a single polypeptide chain comprising, in series, two or
 CC more single domain binding sites. The multivalent antigen proteins
 CC are useful in applications where antibodies have been used in the
 CC prior art, including diagnosis, therapy, targeting, immunoassays,
 CC cross-linking methods including agglutination, or for purification
 CC processes. It is also useful for inactivation of (bacteriophages
 CC or viruses, and in detergents. The present sequence encodes the
 CC heavy chain variable domain of an anti-RR6 antibody from a llama.
 XX
 SQ Sequence 384 BP; 78 A; 96 C; 124 G; 86 T; 0 other;

Query Match 53.6%; Score 299.2; DB 20; Length 384;
 Best Local Similarity 86.2%; Pred. No. 2.5e-72;
 Matches 331; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 12 caggtgcagctcaggaagtcaggggagattgtgacggctgggggtctctgagactc 71
 |||||
 Db 1 caggtgcagctcaggaagtcaggggagattgtgacggctgggggtctctgagactc 60
 |||||
 Qy 72 tctctgacgctcgggacgcgcaccagtggtcatgtgctactatggtatgggtggttc 131

Db 61 tctctgacgctcgggacgcgcacctctcatggtgatggtgctatgcatggctggttc 120
 |||||
 Qy 132 cgcaggttccaggaagagcgtgagtttctgcagctattagtgagtgtaagag 191
 |||||
 Db 121 cgcacaaattccaggaagagcgtgagtttctgcagcaattagtgagtcgtaat 180
 |||||
 Qy 192 acatggtataaagactcctgtaaggccgattccaccatctccagagataacgccaaact 251
 |||||
 Db 181 acatactatgcagactcctgtaaggccgattccaccatctccagagataacgccaaact 240
 |||||
 Qy 252 acggtttatctgcaaatgaacagcctgaacctgaagatacggcgtttattattgtgcc 311
 |||||
 Db 241 atgctgatctgcaaatgaacagttgaaacctgaggacacgcgcgtttactactgtgca 300
 |||||
 Qy 312 gctcagacggctccgctggtgatatttccctgcccgttggtttgactactcggggccag 371
 |||||
 Db 301 gttcggacggtcccggtggttgacatttccagtcgcggttgggttgcctactcggggccag 360
 |||||
 Qy 372 gggaccagggtcacgctctcctca 395
 |||||
 Db 361 gggaccagggtcacgctctcctca 384
 |||||

RESULT 5
 AAX59612
 ID AAX59612 standard; DNA; 714 BP.
 XX
 AC AAX59612;
 XX
 DT 21-JUL-1999 (first entry)
 XX
 DE Plasmid pUR4618 fragment encoding an anti-hCG anti-RR6 ABP.
 XX
 KW Multivalent antigen binding protein; single polypeptide chain;
 KW single domain binding site; diagnosis; therapy; targeting;
 KW immunoassay; cross-linking; agglutination; purification;
 KW phage inactivation; detergent; ss.
 XX
 OS Synthetic.
 XX
 FN WO9923221-A2.
 XX
 PD 14-MAY-1999.
 XX
 PF 27-OCT-1998; 98WO-EP06991.
 XX
 PR 27-OCT-1997; 97EP-0308538.
 XX
 PA (UNIL) UNILEVER NV.
 PA (UNIL) UNILEVER PLC.
 XX
 PI Frenken LGJ, Howell S, Ledebøer AM, Van Der Logt CPE;
 XX
 DR WPI; 1999-313342/26.
 DR
 PT Multivalent antigen binding protein useful for inactivation of
 PT (bacteriophages or viruses
 XX
 PS Disclosure; Fig 27; 89pp; English.
 XX
 CC The specification describes multivalent antigen binding proteins
 CC comprising a single polypeptide chain comprising, in series, two or
 CC more single domain binding sites. The multivalent antigen proteins
 CC are useful in applications where antibodies have been used in the
 CC prior art, including diagnosis, therapy, targeting, immunoassays,
 CC cross-linking methods including agglutination, or for purification
 CC processes. It is also useful for inactivation of (bacteriophages
 CC or viruses, and in detergents. The present sequence encodes an
 CC anti-hCG anti-RR6 bispecific biheaded antigen binding protein (ABP).
 XX
 SQ Sequence 714 BP; 149 A; 181 C; 237 G; 147 T; 0 other;

Query Match	52.2%;	Score 291.2;	DB 20;	Length 714;
Best Local Similarity	85.9%;	Pred. No. 4.5e-70;		
Matches 323;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps
QY	11	acaggtgcagctcagaggtcaggggagagattggtcgagcgtgggggctctctgagact	70	
Db	339	acaggtgcagctcagaggtcaggggagagattggtcgagcgtgggggactctgagact	398	
QY	71	ctctgtgcagctcgggagcgcgccacagtggtgtaattggtcactatggtatgggctgggt	130	
Db	399	ctctgtcgggctcgggacgcactctctatgggtatggtggtatggcatgggctgggt	458	
QY	131	cgccaggttccagggaagagcgtagttgtcgcgagctattaggtggatgggttaaaga	190	
Db	459	cgccaaattccagggaagagcgtagctgtgcgcagcaattaggtggagcgtcgtaa	518	
QY	191	gacatgggtataaagactcogtgaaggcgcgattccacatctccagagataaagcgaagac	250	
Db	519	tacatactatgcagactcogtgaaggcgcgattccacatctccagagataaagcgaagac	578	
QY	251	tacgggttatctgcaaatgaacagcgctgaacctgaagtacagcgcggtttattattgtgc	310	
Db	579	catgctgatctgcaaatgaacagtttgaacctgaggaacgcggtttacactgtgtgc	638	
QY	311	cgctgcacgggtccgggtggatgatatttccctgcgggttgggtttgactactggggcca	370	
Db	639	agttcgagcgttcgcggtggttgacatttccagctcgggttgggtttgcctactggggcca	698	
QY	371	ggggaccaggtcacc	386	
Db	699	qgagaccaggtcacc	714	

RESULT	6	
AAAX59614		
ID	AAAX59614	standard; DNA; 717 BP.
XX	XX	
AAAX59614;		
XX	XX	
21-JUL-1999	(first entry)	
XX	XX	
DE	DE	
XX	XX	Plasmid pUR4620 fragment encoding an anti-hCG anti-RR6 ABP.
XX	XX	
KW	KW	Multivalent antigen binding protein; single polypeptide chain;
KW	KW	single domain binding site; diagnosis; therapy; targeting;
KW	KW	immunoassay; cross-linking; agglutination; purification;
XX	XX	phage inactivation; detergent; ss.
XX	XX	
Synthetic.		
OS	OS	
XX	XX	
WO9923221-A2.		
PN	PN	
XX	XX	
14-MAY-1999.		
PD	PD	
XX	XX	
27-OCT-1998;	98WO-EP06991.	
XX	XX	
27-OCT-1997;	97EP-0308538.	
XX	XX	
(UNIL) UNILEVER NV.		
PA	PA	
(UNIL) UNILEVER PLC.		
XX	XX	
PI	PI	
XX	XX	Frenken LGJ, Howell S, Ledeboer AM, Van Der Logt CPE;
DR	DR	
WPI; 1999-313342/26.		
P-PSDB; AAY15408.		
XX	XX	
PT	PT	Multivalent antigen binding protein useful for inactivation of
PT	PT	(bacterio)phages or viruses
XX	XX	
PS	PS	Disclosure; Fig 10; 89pp; English.
XX	XX	
CC	CC	The specification describes multivalent antigen binding proteins

CC	comprising a single polypeptide chain comprising, in series, two or more single domain binding sites. The multivalent antigen proteins are useful in applications where antibodies have been used in the prior art, including diagnosis, therapy, targeting, immunoassays, cross-linking methods including agglutination, or for purification processes. It is also useful for inactivation of (bacterio)phages or viruses, and in detergents. The present sequence encodes an anti-RR6 anti-RR6 bispecific bishaded antigen binding protein (ABP).					
XX						
SQ	Sequence 717 BP; 158 A; 183 C; 220 G; 156 T; 0 other;					
<hr/>						
Query Match	52.2%;	Score 291.2;	DB 20;	Length 717;		
Best Local Similarity	85.9%;	Pred. No. 4.5e-70;				
Matches 323;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps		
<hr/>						
QY	11	acagggtcagctgcaggagtcagggggagattggtgcaggctgggggctctctcgagact	70			
DB	342	acagggtcagctgcaggagtcagggggagattggtgcaggctgggggctctctcgagact	401			
QY	71	ctctgtgcagctccggagcgcccaccagtggtgatgctcaactatgggtatgggttgggt	130			
DB	402	ctctgtgcagctccggagcgcccaccagtggtgatggttatggtgcatgggttgggt	461			
QY	131	cgcgccagttccagggaagagcgtagttgtcgcagctattagggtgagtggttaaaga	190			
DB	462	ccgcacaattccagggaagagcgtagttgtcgcagctattagggtgagtggttaaaga	521			
QY	191	gacatggtataaagactccgtgaagggccgattccacctcacatctccagagataaacgccaagac	250			
DB	522	tacatactaacyagactccgtgaagggccgattccacctcacatctccagagataaacgccaagac	581			
QY	251	tacgggtttatctgcaaatgaaacagcctgaaacctgaagtacagcgcogtttatattgtgc	310			
DB	582	catgctgtatctgcaaatgaaacagcctgaaacctgaagtacagcgcogtttatattgtgc	641			
QY	311	cgtcgcacgggtccgcgtgatatatttccctgcgggttggtttgactactggggcca	370			
DB	642	agttcggacggtccgcgtgattgacatttccagtcgcgggttggtttgactactggggcca	701			
QY	371	ggggagccccaggtcaccc 386				
DB	702	ggggagccccaggtcaccc 717				
<hr/>						
RESULT	7					
AAx59616						
ID	AAx59616 standard; DNA; 750 BP.					
XX						
AC	AAx59616;					
XX						
DT	21-JUL-1999 (first entry)					
XX						
DE	Plasmid pUR4622 fragment encoding an anti-RR6 antigen binding protein					
XX						
KW	Multivalent antigen binding protein; single polypeptide chain;					
KW	single domain binding site; diagnosis; therapy; targeting;					
KW	immunoassay; cross-linking; agglutination; purification;					
KW	phage inactivation; detergent; ss.					
XX						
OS	Synthetic.					
XX						
PN	WO9923221-A2.					
XX						
PD	14-MAY-1999.					
XX						
PF	27-OCT-1998; 98WO-EF06991.					
XX						
PR	27-OCT-1997; 97EP-0308538.					
XX						
PA	(UNIL) UNILEVER NV.					
XX	(UNIL) UNILEVER P.L.C.					
XX						

20-AUG-1999 (first entry)
 Llama anti-carbazole HC-V fragment 24 DNA.
 Analytical device; electrically conductive solid support; immobilisation;
 electroactive; analysis; peptide hormone; steroid; disease marker;
 diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama;
 HCV-24; ss.
 Llama guanicoe glama.
 WO9927356-A1.
 03-JUN-1999.
 23-NOV-1998; 98WO-GB03495.
 21-NOV-1997; 97EP-0309425.
 (UNIL) UNILEVER NV.
 (UNIL) UNILEVER PLC.
 Badley RA, Porter RA;
 WPI; 1999-385228/32.
 P-PSDB; AAY08985.
 Analytical device including electrochemically active compound on
 solid support
 Example 6.2; Page 66; 78pp; English.
 This invention describes a novel component for an analytical device
 comprising an electrically conductive solid support on which a compound
 (I) is immobilized which has an electroactive part (EA) with an
 electrochemical property that can be modulated in a detectable manner by
 binding it to a specific binding partner (II). Compounds used in the
 method of the invention include those with the formula R = hydrogen,
 hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally
 substituted), halo, amido, or amino; optionally one or more positions on
 the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or
 alkoxy (all optionally substituted), acid groups (organic or inorganic),
 halo, amido or amino. Devices that contain the component of the invention
 are used for qualitative and quantitative analysis of e.g. nucleic acid,
 hormones (peptide or steroid), disease markers, diagnostic indicators
 etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices
 can be used with turbid solutions (whole blood, milk, culture liquids
 etc.), unlike conventional colourimetric methods. This sequence encodes a
 llama HC-V fragment, HCV24 which is used in the method of the invention.
 Sequence 375 BP; 83 A; 90 C; 117 G; 85 T; 0 other;

Query Match 41.7%; Score 232.8; DB 20; Length 375;
 Best Local Similarity 77.6%; Pred. No. 3.6e-54;
 Matches 298; Conservative 0; Mismatches 77; Indels 9; Gaps 1;

QY 12 caggtgcagctgcagagtcaggaggagattggtgcagctgggggctctctgagctc 71
 DB 1 caggtgcagctgcagagtcaggaggagattggtgcagctgggggctctctgagctc 60
 QY 72 tctgtgcagctgcaggcgcgcaccagtggtcatggtcactatggtatggctggttc 131
 DB 61 tctgtgcagctgcaggcgcgcacccttcagttt-----tatgcgctgggttgggttc 111
 QY 132 cgccaggttccagggaagagcgtgagtttgcgcagctattagtgagtggttaagag 191
 DB 112 cgccaggttccagggaagagcgtgagtttgcgcagctattagtgagtggttaagag 171
 QY 192 acatggtataagactcgtgaagggcgattccaccatctccagagataaccccaagact 251
 DB 172 acatactatgcagactcgtgaagggcgattccaccatctccagagataaccccaagact 231

QY 252 acggtttatctgcaaatgaacagcctgaacacctgaagatacagcgcgtttattattgtgcc 311
 DB 232 acggtgtatctgcaaatgaatagcctgaacactagacgatacgcgctttattactgcgca 291
 QY 312 gctcagccggtcgcgctgagatgatatcttcctgcgcggttggttgactactggtggccag 371
 DB 292 gtccgtatgcgctatagtggtgattaccgactatgtaggacatagactactggtggccag 351
 QY 372 gggaccaggtcacccgtctctcca 395
 DB 352 gggaccaggtcacccgtctctcca 375
 RESULT 10
 AAX78159
 ID AAX78159 standard; DNA; 378 BP.
 XX
 AC AAX78159;
 DT 20-AUG-1999 (first entry)
 XX
 DE Llama anti-carbazole HC-V fragment 3 DNA.
 XX
 KW Analytical device; electrically conductive solid support; immobilisation;
 electroactive; analysis; peptide hormone; steroid; disease marker;
 KW diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama;
 KW HCV-3; ss.
 XX
 OS Llama guanicoe glama.
 XX
 PN WO9927356-A1.
 XX
 PD 03-JUN-1999.
 XX
 PF 23-NOV-1998; 98WO-GB03495.
 XX
 PR 21-NOV-1997; 97EP-0309425.
 XX
 PA (UNIL) UNILEVER NV.
 PA (UNIL) UNILEVER PLC.
 XX
 PI Badley RA, Porter RA;
 XX
 DR WPI; 1999-385228/32.
 DR P-PSDB; AAY08984.
 XX
 PT Analytical device including electrochemically active compound on
 solid support
 XX
 PS Example 6.2; Page 65; 78pp; English.
 XX
 CC This invention describes a novel component for an analytical device
 CC comprising an electrically conductive solid support on which a compound
 CC (I) is immobilized which has an electroactive part (EA) with an
 CC electrochemical property that can be modulated in a detectable manner by
 CC binding it to a specific binding partner (II). Compounds used in the
 CC method of the invention include those with the formula R = hydrogen,
 CC hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally
 CC substituted), halo, amido, or amino; optionally one or more positions on
 CC the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or
 CC alkoxy (all optionally substituted), acid groups (organic or inorganic),
 CC halo, amido or amino. Devices that contain the component of the invention
 CC are used for qualitative and quantitative analysis of e.g. nucleic acid,
 CC hormones (peptide or steroid), disease markers, diagnostic indicators
 CC etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices
 CC can be used with turbid solutions (whole blood, milk, culture liquids
 CC etc.), unlike conventional colourimetric methods. This sequence encodes a
 CC llama HC-V fragment, HCV3 which is used in the method of the invention.
 XX
 SQ Sequence 378 BP; 89 A; 94 C; 112 G; 83 T; 0 other;
 Query Match 39.4%; Score 219.8; DB 20; Length 378;

	Best Local Similarity	77.0%;	Pred. No. 1.3e-50;	
	Matches	298;	Conservative	0; Mismatches 77; Indels 12; Gaps
QY	12	caggtgcagctgcaggagtcaggggagagatggtgcagctgggggctctctgcagactc	71	
Db	1	caggtgcagctgcaggagtcaggggagagatggtgcagctgggggctctctgcagactc	60	
QY	72	tcctgtgcagctcgggagcgccacacagtggtcatggtcactatggtatggctggtgttc	131	
Db	61	tcctgtgcagctcggagtcacatga-----ctacctattcaacggctggttc	111	
QY	132	cgcaggttccagggaagagcgtgagttgtgcgcagctattaggtygagtggttaag--	189	
Db	112	cgcaggtccaggggaagagcgtgaattgttaggaatgcttgagtgagtggtggtgc	171	
QY	190	-agacatggtataaagactccgfgaaggccgattccactctccagagataaacgcgaag	248	
Db	172	aacagctactacagactccgfgaaggccgatttaccatctccagagacacgcgaag	231	
QY	249	actacggtttatcgaaatgaacagcctgaaacctgaagatacagccgtttattattgt	308	
Db	232	aatatggtttctcgaaatgagcagcctgaaacctgagggacagccgtttattactgt	291	
QY	309	gccgctgcagcgtccgcgtggatgatatttcctcgcggttggtgttgactactgggac	368	
Db	292	gcagcagcaaacctcacgagtagttacagtgatccgaataattattcattactgggc	351	
QY	369	caggggaccaggtcacgctctctca	395	
Db	352	cagggaaccaggtcacgctctctca	378	

RESULT 11	
AA559605	
ID	AA559605 standard; DNA; 354 BP.
XX	
AC	AA559605;
XX	
DT	21-JUL-1999 (first entry)
XX	
DE	PstI/BstEII fragment of plasmid pUR4602.
XX	
KW	Multivalent antigen binding protein; single polypeptide chain;
KW	single domain binding site; diagnosis; therapy; targeting;
KW	immunoassay; cross-linking; agglutination; purification;
KW	phage inactivation; detergent; ss.
XX	
OS	Synthetic.
XX	
PN	WO923221-A2.
XX	
PD	14-MAY-1999.
XX	
PF	27-OCT-1998; 98WO-EF06991.
XX	
PR	27-OCT-1997; 97EP-0308538.
XX	
PA	(UNIL) UNILEVER NV.
PA	(UNIL) UNILEVER PLC.
XX	
PI	Franken LGJ, Howell S, Ledeboer AM, Van Der Logt CPE;
XX	
XX	
DR	WPI: 1999-313342/26.
DR	P-PSDB; AAY15401.
XX	
PT	Multivalent antigen-binding protein useful for inactivation of
PT	(bacteriophage) or viruses
XX	
PS	Disclosure; Fig 5; 89pp; English.
XX	
CC	The specification describes multivalent antigen binding proteins
CC	comprising a single polypeptide chain comprising, in series, two or
CC	more single domain binding sites. The multivalent antigen proteins

CC are useful in applications where antibodies have been used in the
CC prior art, including diagnosis, therapy, targeting, immunoassays,
CC cross-linking methods including agglutination, or for purification
CC processes. It is also useful for inactivation of (bacterio)phages
CC or viruses, and in detergents. The present sequence encodes the
CC heavy chain variable domain of an anti-hCG antibody HI-15 from a llama.
XX

SQ Sequence 354 BP; 84 A; 93 C; 103 G; 74 T; 0 other;

Query Match 37.6%; Score 209.6; DB 20; Length 354;
Best Local Similarity 76.8%; Pred. No. 8e-48;
Matches 295; Conservative 0; Mismatches 59; Indels 30; Gaps

Qy 12 caggtgcagctgcaggagtccaggggagagattggtcagcgctgggggctctctgcagactc 71
Db 1 caggtgcagctgcaggagtctggggagagaattggtcagcctgggggctctctgaacactc 60

Qy 72 tctgtgcagccttcgggaacgcgccaccagtggtcatggtcaactatgatatggctggttcc 131
Db 61 tctgtgcagccttcggacttaacctca-----ctaatacatgacatggctgggttcc 111

Qy 132 cgccaggttccaggaaagagcgtgagtttgtgcagcactattaggtcgagtggtgtaaagag 191
Db 112 cgccaggtctcaggagtgagccctgtaggcgttatagctggagtggtgataaac 171

Qy 192 acatggtataaagactccgtgaaaggccgatccaccatctcccagatgaacgccaaagact 251
Db 172 acatactatgtaagctccgtgaaaggacgattcacatctcccagatgaacgccaaagac 231

Qy 252 acggtttatctgaaaatgaacagccttgaacacctgaagatcacgcccatttatattgtgcc 311
Db 232 acggtgtatctgaaaatgaacagccttgaacacctgaagacacgcccatttatattgtgca 291

Qy 312 gctcgacccttcgcgtgggatgatatttcctctgcggttggtgtttgactactctggggccag 371
Db 292 gtaaaaccggac-----gatggttggtgggactactctggggccag 330

Qy 372 gggaccacagggtcacgtctctccta 395
Db 331 gagaccacagatcacggtctctccta 354

RESULT	12	
AAx59607		
ID	AAx59607	standard; DNA; 369 BP.
XX	XX	
AC	AAx59607;	
XX	XX	
DT	21-JUL-1999	(first entry)
XX	XX	
DE	PstI/BstEII	fragment of plasmid pUR4642.
XX	XX	
KW	Multivalent antigen binding protein;	single polypeptide chain;
KW	single domain binding site;	diagnosis; therapy; targeting;
KW	immunoassay; cross-linking;	agglutination; purification;
KW	phage inactivation; detergent;	ss.
XX	XX	
OS	Synthetic.	
XX	XX	
PN	W09923221-A2.	
XX	XX	
PD	14-MAY-1999.	
XX	XX	
PF	27-OCT-1998;	98WO-EP06991.
XX	XX	
PR	27-OCT-1997;	97EP-0308538.
XX	XX	
PA	(UNIL) UNILEVER NV.	
PA	(UNIL) UNILEVER PLC.	
XX	XX	
PI	Franken LGJ,	Howell S, Ledeboer AM, Van Der Logt CPE;
XX	XX	

Multivalent antigen binding protein; single polypeptide chain; single domain binding site; diagnosis; therapy; targeting; immunoassay; cross-linking; agglutination; purification; phage inactivation; detergent; ss.

(UNIL) UNILEVER NV.
(UNIL) UNILEVER PLC

Frenken LGJ, Howell S, Ledeboer AM, van Der Logt CPE;

DR WPI; 1999-313342/26.
XX P-PSDB; AAY15403.

PT Multivalent antigen binding protein useful for inactivation of
XX (bacterio)phages or viruses

PS Disclosure; Fig 7; 89pp; English.

XX The specification describes multivalent antigen binding proteins
CC comprising a single polypeptide chain comprising, in series, two or
CC more single domain binding sites. The multivalent antigen proteins
CC are useful in applications where antibodies have been used in the
CC prior art, including diagnosis, therapy, targeting, immunoassays,
CC cross-linking methods including agglutination, or for purification
CC processes. It is also useful for inactivation of (bacterio)phages
CC or viruses, and in detergents. The present sequence encodes the
CC heavy chain variable domain of an anti-Streptococcus antibody S120
CC from a llama.

XX Sequence 369 BP; 88 A; 90 C; 109 G; 82 T; 0 other;

Query Match 37.5%; Score 209.2; DB 20; Length 369;

Best Local Similarity 75.8%; Pred. No. 1e-47; Mismatches 15; Gaps 2;

Matches 291; Conservative 0; Indels 15; Gaps 2;

Qy 12 cagggtcagctcagagtcagggtgaggtggtgaggtggtgaggtggtgaggtc 71

Db 1 cagggtcagctcagagtcagggtgaggtggtgaggtggtgaggtggtgaggtc 60

Qy 72 tctgtcagctcagagtcagggtgaggtggtgaggtggtgaggtggtgaggtc 131

Db 61 tctgtcagctcagagtcagggtgaggtggtgaggtggtgaggtggtgaggtc 111

Qy 132 ggcaggttcagggtgaggtgaggtggtgaggtggtgaggtggtgaggtgag 191

Db 112 ggcaggttcagggtgaggtgaggtggtgaggtggtgaggtggtgaggtgag 171

Qy 192 acatgtataagactccgtgaggtgaggtgaggtgaggtgaggtgaggtgag 251

Db 172 acatgtataagactccgtgaggtgaggtgaggtgaggtgaggtgaggtgag 231

Qy 252 acggtttatctgcaaatgaacagctgaaacctgaagatacggcgtttattgtgc 311

Db 232 acaggttatctcaaatgaacagctgaaacctgaagatacggcgtttattgtgc 291

Qy 312 gctcagccgctcgtgaggtgaggtgaggtgaggtgaggtgaggtgaggtgag 371

Db 292 gccagtaggtattactatgcccacatgccc-----aatcagtagctgactggtggccag 345

Qy 372 gggaccaggtcagctcctca 395

Db 346 gggaccaggtcagctcctca 369

RESULT 13

AAX59604

ID AAX59604 standard; DNA; 351 BP.

XX AAX59604;

XX AAX59604;

XX 21-JUL-1999 (first entry)

XX PstI/BstEII fragment of plasmid pUR4601.

XX Multivalent antigen binding protein; single polypeptide chain;

XX single domain binding site; diagnosis; therapy; targeting;

XX immunoassay; cross-linking; agglutination; purification;

XX phage inactivation; detergent; ss.

XX Synthetic.

XX OS

XX W09923221-A2.

XX

XX

XX

XX

XX

XX

XX

XX

XX 14-MAY-1999.

XX 27-OCT-1998; 98WO-EP06991.

XX 27-OCT-1997; 97EP-0308538.

XX (UNIL) UNILEVER NV.

XX (UNIL) UNILEVER PLC.

XX Frenken LGJ, Howell S, Ledebroer AM, Van Der Logt CPE;

XX WPI; 1999-313342/26.

XX P-PSDB; AAY15403.

XX Multivalent antigen binding protein useful for inactivation of

XX (bacterio)phages or viruses

XX Disclosure; Fig 4; 89pp; English.

XX The specification describes multivalent antigen binding proteins

XX comprising a single polypeptide chain comprising, in series, two or

XX more single domain binding sites. The multivalent antigen proteins

XX are useful in applications where antibodies have been used in the

XX prior art, including diagnosis, therapy, targeting, immunoassays,

XX cross-linking methods including agglutination, or for purification

XX processes. It is also useful for inactivation of (bacterio)phages

XX or viruses, and in detergents. The present sequence encodes the

XX heavy chain variable domain of an anti-hCG antibody H14 from a llama.

XX Sequence 351 BP; 75 A; 91 C; 120 G; 65 T; 0 other;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

linker; conformational flexibility; multivalent binding protein;
 KW bi-head; immunoassay; agglutination assay; purification; antibody;
 KW human chorionic gonadotropin; hCG; ds.
 XX
 OS Lama guanicoe glama.
 OS Homo sapiens.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..351
 FT /*tag= a
 FT /product= "Llama HC-V domain of anti-hCG antibody (H14)"
 FT /note= "Does not include stop codon"
 FT /partial
 XX
 PN WO200024884-A2.
 XX
 XX 04-MAY-2000.
 XX
 XX 22-OCT-1999; 99WO-EP08323.
 XX
 XX 27-OCT-1998; 98WO-EP06991.
 PR 22-APR-1999; 99EP-0303118.
 PR
 XX (UNIL) UNILEVER PLC.
 PA (UNIL) UNILEVER NV.
 PA (HIND-) HINDUSTAN LEVER LTD.
 XX
 XX Frenken LGJ, Howell S, Van Der Vaart JM;
 XX
 XX WPI; 2000-350728/30.
 DR P-PSDB; AAY70992.
 XX
 XX Use of a linker whose amino acid sequence confers restricted
 PT conformational flexibility to generate multivalent and multispecific
 PT antigen binding proteins -
 PT
 XX Example 1.1b; Fig 2; 50pp; English.
 XX
 XX The present DNA sequence is the PstI-BstEII insert of the plasmid
 CC PUR4601, encoding llama heavy chain variable (HC-V) domain of an anti-hCG
 CC antibody (denoted as H14). The PUR4601 plasmid comprising this sequence
 CC is transformed into *Saccharomyces cerevisiae*. They are screened for
 CC antigen specific HC-V fragments that specifically bind to human chorionic
 CC gonadotropin (hCG) antigen. This plasmid is used for the construction of
 CC antigen binding proteins, comprising two single binding units referred to
 CC as Bi-heads with various linkers. The peptide linker confers restricted
 CC conformational flexibility for linking binding units in a multivalent
 CC binding protein. The linker is used to generate multivalent or
 CC multispecific antigen binding proteins for immunoassays, agglutination
 CC assays or for purification.
 XX
 XX Sequence 351 BP; 75 A; 91 C; 120 G; 65 T; 0 other;
 SQ

Query Match 37.3%; Score 208.4; DB 21; Length 351;
 Best Local Similarity 83.4%; Pred. No. 1.7e-47;
 Matches 252; Conservative 0; Mismatches 41; Indels 9; Gaps 1;

QY 12 cagggtcagctcaggagtcaggagagagtggtgagctgggggctctctgagactc 71
 |||||
 DB 1 cagggtcagctcaggagtcaggagagagtggtgagctgggggctctctgagactc 60
 |||||
 QY 72 tccgtgagctcgggagcgccaccagtggtcagctatggtatgggtggttc 131
 |||||
 DB 61 tccgtgagctcgtggagcgaccggca-----gtacgtatgacatgggctggttc 111
 |||||
 QY 132 cggcaggttccagggaagagcgtgagttgtcgcagctattagggtgagtgtaaaagag 191
 |||||
 DB 112 cggcagggccggggaagagcgtgagttgtcgcagctattagggtgagtgtaaaagag 171
 |||||
 QY 192 acatggtataaagactccgtgaggccgctatccaccatctccagagataacgccagact 251
 |||||

Db 172 acatactatgcaagctccgtgagggccgattccaccatctccagagacaaacgcaagaag 231
 QY 252 acggtttatctcaaatgaacagctggaacctgaagatacagccgtttattattgtgcc 311
 |||||
 Db 232 acggtgtatctgcaaatgaacagctggaacctgaagatacagccgtttattattgtgcc 291
 QY 312 gc 313
 ||
 Db 292 gc 293

RESULT 15
 AAD10055
 ID AAD10055 standard; DNA; 462 BP.
 XX
 AC AAD10055;
 XX
 XX 18-SEP-2001 (first entry)
 XX
 XX Anti-GUS (Clone18) VH region attached with myc and his6 tag DNA.
 XX
 KW Heavy chain immunoglobulin; pathogen resistance; metabolism modulator;
 KW passive immunisation; heavy chain variable domain; VH; anti-GUS antibody;
 KW beta-glucuronidase; ds.
 XX
 OS Unidentified.
 XX
 XX Key Location/Qualifiers
 FT CDS 1..453
 FT /*tag= a
 FT /product= "Anti-GUS heavy chain variable domain
 FT with myc and his6 peptide"
 FT /note= "Does not include start codon"
 FT /partial
 XX
 PN EP118669-A2.
 XX
 XX 25-JUL-2001.
 XX
 XX 08-DEC-2000; 2000EP-0310997.
 XX
 XX 17-DEC-1999; 99EP-0310188.
 XX
 XX (UNIL) UNILEVER PLC.
 PA (UNIL) UNILEVER NV.
 XX
 XX Frenken LGJ, Van Der Logt CPE, Jobling SA, Teh Y;
 PI
 XX WPI; 2001-427157/46.
 DR P-PSDB; AAE05284.
 XX
 XX Modifying a plant to produce an antibody useful for increasing pathogen
 PT resistance or to modulate metabolism comprises introducing a DNA
 PT sequence encoding a heavy chain immunoglobulin linked to a peptide that
 PT targets a cellular compartment -
 PT
 XX Example 1; Fig 14; 81pp; English.
 XX
 XX The present invention relates to a method for modifying a plant to
 CC produce an antibody or an active fragment or derivative, or a protein
 CC functional equivalent, in a cellular compartment comprising introducing
 CC a DNA sequence encoding a heavy chain immunoglobulin, where the DNA is
 CC linked to promoters and provided with an additional sequence encoding a
 CC peptide capable of targeting heavy chain immunoglobulin to a cellular
 CC compartment. The method is used for producing a heavy chain
 CC immunoglobulin or an active fragment or derivative, or a protein that is
 CC functionally equivalent for increasing the pathogen resistance in a plant
 CC or to modulate metabolism in a plant. Under some circumstances it may be
 CC desirable to retain the antibody product with the plant rather than
 CC extracting and isolating the product. In particular, edible selected
 CC antigens may be used in a method of passively immunising an animal,
 CC preferably human, against the antigen, e.g., pathogenic organisms. The
 CC present DNA sequence encodes an anti-GUS antibody (denoted Clone18)

CC heavy chain variable domain (VH) attached to peptide linkers, myc and
CC his6 tag. The GUS is beta-glucuronidase.

XX
SQ Sequence 462 BP; 105 A; 136 C; 138 G; 83 T; 0 other;

```
Query Match      37.3%; Score 208; DB 22; Length 462;
Best Local Similarity 72.8%; Pred. NO. 2.3e-47;
Matches 303; Conservative 0; Mismatches 95; Indels 18; Gaps 2;

QY 12 cagggtcagctcaggagtcaggaggagattggtgcaagctgggggctctctgagactc 71
   ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10 cagggtgaaactcagcagctctgaggaggagattggtgcaagctgggggctctctgagctc 69
   ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 72 tctgtgacgtcggagcgcgccaccagtggtcactatggtatggtgctggttc 131
   ||||| ||||| ||||| || || || || || || || || || || || || || || || ||
Db 70 tctgtgacgtcctggacgcacctca-----gtaactatgccgtgggctggttc 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 132 cgccagggtccaggaggagagcgtgagttgtcgagctattagtgagtggtgtaagag 191
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 cgccagggtccaggaggagcgtgagttgtcgctgctattagccgtgattggtgggcg 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 192 acatggtataaagactccgtgaagcgcgattccaccatctccagagataacgccaaagact 251
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 acatactatgcggactccgtggaaggccgattccgctctccagagactcacgccagagaac 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 252 acggtttatctgcaaatgaacagcctgaaacctgaagatacggccggtttattatgtgoc 311
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 acggtgtatctgcaaatgaacagcctgaaacctgaagatacggccggtttattatgtgtaac 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 312 gctgaccggtccgcgtggatgatatttccctggccggttggtgttgactactggggccag.371
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 acaaggccctactggggccagggggacccagggtcacct-----ctcctcagcgcac 351
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 372 gggaccagggtcacctctcctcaggatctcatcaccatcacccatcaggatccac 427
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 352 cacagcgaagacccagctcccgccgcccatcaccatcacccatcacggggccgc 407
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: August 13, 2002, 12:36:32
Job time: 11545 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 06:47:17 ; Search time 1846.07 Seconds
(without alignments)
6325.333 Million cell updates/sec

Title: US-09-742-690-1
Perfect score: 558
Sequence: 1 tcgagaagaacaggtgcag.....ctcccagtgctgtataaag 558

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba:

2: gb_hgt:

3: gb_in:

4: gb_om:

5: gb_ov:

6: gb_pat:

7: gb_ph:

8: gb_pl:

9: gb_pr:

10: gb_ro:

11: gb_sts:

12: gb_sy:

13: gb_un:

14: gb_vi:

15: em_ba:

16: em_fun:

17: em_hum:

18: em_in:

19: em_mu:

20: em_om:

21: em_or:

22: em_ov:

23: em_pat:

24: em_ph:

25: em_pl:

26: em_ro:

27: em_sts:

28: em_un:

29: em_vi:

30: em_htg_hum:

31: em_htg_inv:

32: em_htg_other:

33: em_htgo_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1

AX164056

LOCUS

AX164056

DEFINITION

AX164056

ACCESSION

AX164056.1

VERSION

GI:14545010

KEYWORDS

synthetic construct.

ORGANISM

synthetic construct

REFERENCE

1 (bases 1 to 440)

AUTHORS

Frenken, L.G., van der Logt, C.P., Jobling, S.A. and Teh, Y.M.

TITLE

Production of camelid antibodies in plants

JOURNAL

Patent: EP 1118669-A 1 25-JUL-2001;

UNILEVER PLC (GB); UNILEVER N.V. (NL)

FEATURES

Location/Qualifiers

source

1..440

AX164056 Sequence

AX180311 Sequence

AX164109 Sequence

AX164111 Sequence

AJ236100 Lama glam

AJ236108 Lama glam

AJ236106 Lama glam

AJ236099 Lama glam

AJ236103 Lama glam

AJ236104 Lama glam

AJ236105 Lama glam

AX021690 Sequence

AX021717 Sequence

AX021721 Sequence

AX021725 Sequence

AX021727 Sequence

AF442920 Lama glam

AF441487 Lama glam

AF442907 Lama glam

AF442908 Lama glam

AF442923 Lama glam

AJ238059 Lama glam

A95262 Sequence 8

AF442927 Lama glam

AF442915 Lama glam

AX164056 440 bp

Sequence 1 from Patent EP1118669.

AX164056

AX164056.1

GI:14545010

synthetic construct.

synthetic construct

artificial sequence.

1 (bases 1 to 440)

Frenken, L.G., van der Logt, C.P., Jobling, S.A. and Teh, Y.M.

Production of camelid antibodies in plants

Patent: EP 1118669-A 1 25-JUL-2001;

UNILEVER PLC (GB); UNILEVER N.V. (NL)

Location/Qualifiers

1..440

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="VHH with peptide linker"

<1..420

/note="unnamed protein product"

/codon_start=1

/transl_table=11
 /protein_id="CAC42706.1"
 /db_xref="GI:14545011"
 /translation="OVOLQESGGGLVOAGSLRLSCAASGRATSGHGHYGMWPRQVP
 GKREFFVAALRWSEKETYKDSYKGRFTISRDNKATTVYIOMNSLKRPEDTAVYYCAAR
 PVRVDDISLPVDFYQGGTQVTVSSQKLISEEDLN"
 BASE COUNT 103 a 103 c 135 g 99 t
 ORIGIN

Query Match 69.1%; Score 385.4; DB 6; Length 440;
 Best Local Similarity 99.7%; Pred. No. 7.2e-111;
 Matches 386; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 12 caggtgcagctcaggaagtcaggggagagattggtgcaggctgggggctctctgagactc 71
 |||||

Db 1 CAGGTGCAGCTGCAGGACTCAGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTC 60
 |||||

Qy 72 tccgtgcagctcaggaagtcaggggagagattggtgcaggctgggggctctctgagactc 131
 |||||

Db 61 TCCTGTGCAGCTGCAGGACTCAGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTC 120
 |||||

Qy 132 cgcaggttcaggaagtcaggggagagattggtgcaggctgggggctctctgagactc 191
 |||||

Db 121 GCCCAGGTTCCAGGAGGAGCGTGAGTTTGTGCGAGCTATTAGGTGAGTGGTAAAGAG 180
 |||||

Qy 192 acatgtataaagactcogtgaaggccgattccaccatctccagagataaagcgaagact 251
 |||||

Db 181 ACATGGTATAAAGACTCCGTGAAGGCCGATTCACCATCTCCAGAGATAACGCCAAGACT 240
 |||||

Qy 252 acggtttatctgcaaatgaacagcctgaacctgaagatacggcgctttattattgtgcc 311
 |||||

Db 241 ACGGTTTATCTGCAATGAACAGCCTGAACCTGAAGATACGGCGCTTTATTATTGTGCC 300
 |||||

Qy 312 gctcagccggtcccgctgatgatattccctgcgggttggttgactactgagggcag 371
 |||||

Db 301 GCTGACCGGTCGCGTGAGATGATATTTCCCTGCGGTTGGGTTTGACTACTGGGGCCAG 360
 |||||

Qy 372 gggaccaggtcacgctcctcagatctcatcaccatccacatcacgga 422
 |||||

Db 361 GGGACCCAGGTACCGTCTCTCACAGGTGCAGCTGCAGCAGCTCTGGGGGA 411
 |||||

RESULT 2
 AX180311
 LOCUS AX180311 888 bp DNA linear PAT 06-AUG-2001

DEFINITION Sequence 9 from Patent WO0146364.
 ACCESSION AX180311
 VERSION AX180311.1 GI:15132274

KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 888)
 AUTHORS Howell, S., Little, J., van der Logt, C.P. and Parry, N.J.
 TITLE Method of delivering a benefit agent
 JOURNAL Patent: WO 0146364-A 9 28-JUN-2001;
 FEATURES Location/Qualifiers
 1..888
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="anti-RR6-VHH8-his-CBD"
 BASE COUNT 200 a 230 c 254 g 204 t
 ORIGIN

Query Match 69.1%; Score 385.4; DB 6; Length 888;
 Best Local Similarity 96.1%; Pred. No. 7.7e-111;
 Matches 395; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy 12 caggtgcagctcaggaagtcaggggagagattggtgcaggctgggggctctctgagactc 71
 |||||

Db 1 CAGGTGCAGCTGCAGGAGTACGGGGAGGATTGGTGCAGCTGGGGGCTCTCTGAGACTC 60
 |||||

Qy 72 tccgtgcagctcaggaagtcaggggagagattggtgcaggctgggggctctctgagactc 131
 |||||

Db 61 TCCTGTGCAGCTGCAGGACTCAGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTC 120
 |||||

Qy 132 cgcaggttcaggaagtcaggggagagattggtgcaggctgggggctctctgagactc 191
 |||||

Db 121 GCCCAGGTTCCAGGAGGAGCGTGAGTTTGTGCGAGCTATTAGGTGAGTGGTAAAGAG 180
 |||||

Qy 192 acatgtataaagactcogtgaaggccgattccaccatctccagagataaagcgaagact 251
 |||||

Db 181 ACATGGTATAAAGACTCCGTGAAGGCCGATTCACCATCTCCAGAGATAACGCCAAGACT 240
 |||||

Qy 252 acggtttatctgcaaatgaacagcctgaacctgaagatacggcgctttattattgtgcc 311
 |||||

Db 241 ACGGTTTATCTGCAATGAACAGCCTGAACCTGAAGATACGGCGCTTTATTATTGTGCC 300
 |||||

Qy 312 gctcagccggtcccgctgatgatattccctgcgggttggttgactactgagggcag 371
 |||||

Db 301 GCTGACCGGTCGCGTGAGATGATATTTCCCTGCGGTTGGGTTTGACTACTGGGGCCAG 360
 |||||

Qy 372 gggaccaggtcacgctcctcagatctcatcaccatccacatcacgga 422
 |||||

Db 361 GGGACCCAGGTACCGTCTCTCACAGGTGCAGCTGCAGCAGCTCTGGGGGA 411
 |||||

RESULT 3
 AX164109
 LOCUS AX164109 1154 bp DNA linear PAT 22-JUN-2001

DEFINITION Sequence 54 from Patent EP1118669.
 ACCESSION AX164109
 VERSION AX164109.1 GI:14545059

KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 1154)
 AUTHORS Franken, L.G., van der Logt, C.P., Jobling, S.A. and Teh, Y.M.
 TITLE Production of camelid antibodies in plants
 JOURNAL Patent: EP 1118669-A 54 25-JUL-2001;
 UNILEVER PLC (GB); UNILEVER N.V. (NL)

FEATURES Location/Qualifiers
 1..1154
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="HCV33-hinge-CH2-CH3"
 3..1136
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC42712.1"
 /db_xref="GI:14545060"
 /translation="MEVQLQESGGGLVQAGSLRLSCAASGRATSGHGHYGMWPRQV
 PGRKRFVAALRWSEKETYKDSYKGRFTISRDNKATTVYIOMNSLKRPEDTAVYYCAA
 RPVRVDDISLPVDFYQGGTQVTVSSQKLISEEDLN"
 BASE COUNT 293 a 341 c 321 g 199 t
 ORIGIN

Query Match 69.0%; Score 384.8; DB 6; Length 1154;
 Best Local Similarity 99.5%; Pred. No. 1.2e-110;
 Matches 386; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 13 aggtgcagctcaggaagtcaggggagagattggtgcaggctgggggctctctgagactc 72
 |||||

Db 7 AGGTGCACCTCAGCAGTACAGGGGAGGATTGGTGCAGCTGGGGGCTCTCTGAGACTCT 66
 |||||

Qy 73 cctgtgcagctcaggaagtcaggggagagattggtgcaggctgggggctctctgagactc 132
 |||||


```

Db 67 CCGTGTGAGCCTCGGAGCGCCGCCACAGTGGTGCATGTCACACTATGTATGGGCTGGTTC 126
QY 133 gccaggtccaggaagagcgtagtttgcagctattagtgagtgagtgtaagaga 192
|||||
Db 127 GCCAGGTTCCAGGAGGAGCGTAGTTGTCCACACTATAGTGGAGTGGTAAAGAGA 186
QY 193 catggtataaagactccgtgaagggcgattccaccatctccagagataacgccaaagacta 252
|||||
Db 187 CATGGTATAAAGACTCCGTGAAGGCGGATTCACCATCTCCAGAGATAACGCCCAAGACTA 246
QY 253 cgtttatctgaaatgaacagcctgaacactgaagatacagcggtttattattgtgcg 312
|||||
Db 247 CGGTTTATCTGCAATGAACAGCCTGAACCTGAAGATAGCGCGGTTATTATTGTGCCG 306
QY 313 ctgacccggtcccggtgagatattccctgcgcggttggttgactactgagggccagg 372
|||||
Db 307 CTCGACCGGTCCCGGTGATGATATTCCCTGCGGTTGGGTTTGACTACTGGGGCCAGG 366
QY 373 ggaccaggtcacccgtctctcctcaggatc 400
|||||
Db 367 GGACCCAGGTCACCGTCTCTCAGAACCC 394

RESULT 4
AX164111
LOCUS AX164111 1172 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 56 from Patent EP1118669.
ACCESSION AX164111
VERSION AX164111.1 GI:14545061
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 1172)
AUTHORS Frenken,L.G., van der Loof,C.P., Jobling,S.A. and Teh,Y.M.
TITLE Production of camelid antibodies in plants
JOURNAL Patent: EP 1118669-A 56 25-JUL-2001;
UNILEVER PLC (GB) ; UNILEVER N.V. (NL)
FEATURES
Source 1..1172
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="hinge-HCV33-CH2-CH3-SEKDEL"
CDS 3..1154
/notes="unnamed protein product"
/codon_start=1
/trans_cable=11
/db_protein_id="CAC42713.1"
/db_xref="GI:14545062"
/translation="MEVQLQESGGGLVQAGSLRLSCAASGRATSGHGHGMGFPROV
PKREFVAAIRWSGKETWYKDSVKGRFTISRDNKATTVYLOMNSLKPEDTAVYYCAA
RPVRVDDISLPVGFYWGCGTQVTVSSSEKTPKPPQPPQPPQPPQPPPTTESKPKCPAP
ELGGSPVFIFPKPRVDVLSISGRPEVTCVVDVGQEDPEVSNFTYIDGAERTANTR
PKPEQFNSTYVPLVPIQIHDMLTGKEFKKNNKALPAIEKTSKAKGOTREPOV
YALAHRELDKDFVSTCLVKGFPDPIINVEWRNQPSSEGTATTPOLDNDGTY
FLYSLSVCKNTWQRCETFTCVVMEALHNHTQKSTQSSGKSEKDEL"
BASE COUNT 300 a 343 c 326 g 203 t
ORIGIN

Query Match 69.0%; Score 384.8; DB 6; Length 1172;
Best Local Similarity 99.5%; Pred. No. 1-2e-110;
Matches 386; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 aggtgcagctgcaggagtcaggggaggattggtgcaggtcggggctctctgagactct 72
|||||
Db 7 AGGTGCAGCTGCAGGAGTCAAGGGGAGGATTGGTCAGGCTGGGGGCTCTCAGACTCT 66
QY 73 cctgtgcagctcgggagcgagcaccagtggttcagtgctactatggtggtggttc 132
|||||
Db 67 CCGTGTGAGCCTCGGAGCGCCGCCACAGTGGTGCATGTCACACTATGTATGGGCTGGTTC 126
QY 133 gccaggttcaggaagagcgtagtttgcagcagctattagtgagtgagtgtaagaga 192

```

```

|||||
Db 127 GCCAGGTTCCAGGAGGAGCGTAGTTGTCCAGCTATTAGTGGAGTGGTAAAGAGA 186
QY 193 catggtataaagactccgtgaagggcgattccaccatctccagagataacgccaaagacta 252
|||||
Db 187 CATGGTATAAAGACTCCGTGAAGGCGGATTCACCATCTCCAGAGATAACGCCCAAGACTA 246
QY 253 cgtttatctgaaatgaacagcctgaacactgaagatacagcggtttattattgtgcg 312
|||||
Db 247 CGGTTTATCTGCAATGAACAGCCTGAACCTGAAGATAGCGCGGTTATTATTGTGCCG 306
QY 313 ctgacccggtcccggtgagatattccctgcgcggttggttgactactgagggccagg 372
|||||
Db 307 CTCGACCGGTCCCGGTGATGATATTCCCTGCGGTTGGGTTTGACTACTGGGGCCAGG 366
QY 373 ggaccaggtcacccgtctctcctcaggatc 400
|||||
Db 367 GGACCCAGGTCACCGTCTCTCAGAACCC 394

RESULT 5
LGL236100
LOCUS LGL236100 384 bp DNA linear MAM 09-OCT-2000
DEFINITION Lama glama mRNA for Ig heavy chain variable region (r2).
ACCESSION AJ236100
VERSION AJ236100.1 GI:4165531
KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable
region.
SOURCE llama.
ORGANISM Lama glama
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
REFERENCE 1 (bases 1 to 384)
AUTHORS van der Linden,R.H., de Geus,B., Frenken,G.J., Peters,H. and
Verrips,C.T.
TITLE Improved production and function of llama heavy chain antibody
fragments by molecular evolution
JOURNAL J. Biotechnol. 80 (3), 261-270 (2000)
MEDLINE 20403592
REFERENCE 2 (bases 1 to 384)
AUTHORS Frenken,L.G.J.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever
Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE
NETHERLANDS
FEATURES
Source 1..384
/organism="Lama glama"
/db_xref="taxon:9844"
/cell_type="peripheral blood lymphocyte"
/rearranged
V_region <1..>384
/gene="r2"
gene 1..384
/gene="r2"
CDS <1..>384
/gene="r2"
/codon_start=1
/protein_id="CAA15412.1"
/db_xref="GI:4165532"
/translation="QVQLQESGGGLVQAGSLRLSCAASGRATSGHGHGMGFPROVP
GKEREFVAAIRWSGKETWYKDSVKGRFTISRDNKATTVYLOMNSLKPEDTAVYYCAAR
PYRVDDISLPVGFYWGCGTQVTVSS"
BASE COUNT 80 a 93 c 124 g 87 t
ORIGIN

Query Match 68.8%; Score 384; DB 4; Length 384;
Best Local Similarity 100.0%; Pred. No. 2e-110;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 caggtgcagctgcaggagtcaggggaggattggtgcaggtcggggctctctgagactc 71
|||||

```

```

Db 1 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTCAGAGCTGGGGCTCTCTGAGACTC 60
QY 72 tctgtgagcctcgggacgcgcacacagtggtcatgtgctcaactatggtatggcgctgttc 131
Db 61 TCCTGTGAGCCTCGGACGCGCCACCAGTGGTCATGGTCACATATGCTATGGGTGGCTGCTC 120
QY 132 cgcaggttcagggaagagcgtgagttgtgcagcactattagtgagtggttaagag 191
Db 121 CGCAGAGTTCCAGGGAAGGAGCGTGAGTTTGTGCGACGTATTAGTGGAGTGGTAAAGAG 180
QY 192 acatggtataaagactccgtgaagggccgattccaccatctccagagataacgcgaagact 251
Db 181 ACATGTTATAAGACTCCGTGAAGGGCGGATTACCATCTCCAGAGATACGCCAAAGACT 240
QY 252 acggtttatctgcaaatgaacagcctgaaacactgaaagatacgcgcgtttattattgtgcc 311
Db 241 ACAGTTTATCTGCAAAATGAACAGCCTGAAACCTGAAGATACGGCGGTTTATTATTGTGCC 300
QY 312 gctcgaccggtccgcgtgagtgatatttccctgcgcgttggttgactactggggccag 371
Db 301 GCTCGACCGGTCCGCGTGGATGATATTTCCTCCGCGTGGGTTTGACTACTGGGGCCAG 360
QY 372 gggaccaggtccacgctctctctca 395
Db 361 GGGACCCAGGTCACCGTCTCTCTCA 384

RESULT 6
LGL236108
LOCUS
DEFINITION
Lama glama mRNA for Ig heavy chain variable region (r10).
ACCESSION
AJ236108.1 GI:4165547
VERSION
immunoglobulin heavy chain; immunoglobulin superfamily; variable
region.
SOURCE
llama.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
REFERENCE
1 (bases 1 to 384)
AUTHORS
van der Linden,R.H., de Geus,B., Frenken,G.J., Peters,H. and
Verrips,C.T.
TITLE
Improved production and function of llama heavy chain antibody
fragments by molecular evolution
JOURNAL
J. Biotechnol. 80 (3), 261-270 (2000)
MEDLINE
20403592
REFERENCE
2 (bases 1 to 384)
AUTHORS
Frenken,L.G.J.
TITLE
Direct Submission
Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever
Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE
NETHERLANDS
FEATURES
Location/Qualifiers
source
1..384
/organism="Lama glama"
/db_xref="taxon:9844"
/cell_type="peripheral blood lymphocyte"
/rearranged
<1..>384
/gene="r10"
/codon_start=1
/protein_id="CAA15420.1"
/db_xref="GI:4165548"
/translation="QVQLVSGGGLVQAGSGLRSLCAASGRASSCHGGYGMGWFRRQVP
GKREFAAIRWSGKETWYDVSVKGRFTISRDNKSTVYLQMSLKPEDTAVYICAVR
PVRVDDISSPVGFDYWGQGTQVTVSS"
BASE COUNT 78 a 94 c 129 g 83 t
ORIGIN

```

```

Query Match 64.5%; Score 360; DB 4; Length 384;
Best Local Similarity 96.1%; Pred. No. 8e-103;
Matches 369; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 12 caagtcacagtcagagatcaggggagagattggtcagagctaggggctctctgagactc 71
Db 1 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTCAGAGCTGGGGCTCTCTGAGACTC 60
QY 72 tctgtgagcctcgggacgcgcacacagtggtcagcactatggtatggcgctgttc 131
Db 61 TCCTGTGAGCCTCGGACGCGCCACCAGTGGTCATGGTCACATATGCTATGGGTGGCTGCTC 120
QY 132 cgcaggttcagggaagagcgtgagttgtgcagcactattagtgagtggttaagag 191
Db 121 CGCAGAGTTCCAGGGAAGGAGCGTGAGTTTGTGCGACGTATTAGTGGAGTGGTAAAGAG 180
QY 192 acatggtataaagactccgtgaagggccgattccaccatctccagagataacgcgaagact 251
Db 181 ACATGTTATAAGACTCCGTGAAGGGCGGATTACCATCTCCAGAGACACGCCAAAGACT 240
QY 252 acggtttatctgcaaatgaacagcctgaaacactgaaagatacgcgcgtttattattgtgcc 311
Db 241 ACAGTTTATCTGCAAAATGAACAGCCTGAAACCTGAAGATACGGCGGTTTATTATTGTGCC 300
QY 312 gctcgaccggtccgcgtgagtgatatttccctgcgcgttggttgactactggggccag 371
Db 301 GTCGACCGGTCCGCGTGGATGATATTTCAGGCCGCGTGGGTTTGACTACTGGGGCCAG 360
QY 372 gggaccaggtccacgctctctctca 395
Db 361 GGGACCCAGGTCACCGTCTCTCTCA 384

RESULT 7
LGL236106
LOCUS
DEFINITION
Lama glama mRNA for Ig heavy chain variable region (r8).
ACCESSION
AJ236106
VERSION
AJ236106.1 GI:4165543
KEYWORDS
immunoglobulin heavy chain; immunoglobulin superfamily; variable
region.
SOURCE
llama.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
REFERENCE
1 (bases 1 to 384)
AUTHORS
van der Linden,R.H., de Geus,B., Frenken,G.J., Peters,H. and
Verrips,C.T.
TITLE
Improved production and function of llama heavy chain antibody
fragments by molecular evolution
JOURNAL
J. Biotechnol. 80 (3), 261-270 (2000)
MEDLINE
20403592
REFERENCE
2 (bases 1 to 384)
AUTHORS
Frenken,L.G.J.
TITLE
Direct Submission
Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever
Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE
NETHERLANDS
FEATURES
Location/Qualifiers
source
1..384
/organism="Lama glama"
/db_xref="taxon:9844"
/cell_type="peripheral blood lymphocyte"
/rearranged
<1..>384
/gene="r8"
/codon_start=1
/protein_id="CAA15418.1"
V_region
gene
1..384
/gene="r8"
<1..>384
/gene="r8"
CDS

```

/db_xref="GI:4165544"
/translation="QVQLQSGGGLVQAGSLRLSCVASGRTFSCHGYGNGWFRQVP
GKERELVAARIWMSGLETHKDSVGRFTISRDNKSTVYLQNSLKPEDTGYVYCAVR
PVRVDDISPPGVFGQGTQVTVSS"
BASE COUNT 76 a 94 c 126 g 88 t
ORIGIN

Query Match 58.5%; Score 326.4; DB 4; Length 384;
Best Local Similarity 90.6%; Pred. No. 3.5e-92;
Matches 348; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 12 cagggtgcagctcagagagtcaggggagagattggtgcagctcgggggctctctgagactc 71
DB 1 CAGGTGCAGCTCAGCAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCTGAGACTC 60
QY 72 tctgtgcagctcgggacgcgccaccagtggtcagctcagctcagctcagctcagctc 131
DB 61 TCTGTGTAGCTTCGGGACGCGACCTTCAGTGGCCATGGTGTGCGCATGGCTGGCTGTTTC 120
QY 132 cgccaggttccagggaagagcgtgagttgtcgcagctcattaggtggagtggttaagag 191
DB 121 CGCCAGGTTCCAGGGAAGCAGCGTGAGTTGTGCGCAGCTATTAGGTGGAGCGGTGTACG 180
QY 192 acatggtataaagactccgtgaaggccagattccatctccagagataaacgccaagact 251
DB 181 ACATATTATGTAGTACCTCCGTGAAGGGCCGATTCCACCATCTCCAGAGACAAACGCAAGAGC 240
QY 252 acggtttatctcaaatgaacagcctgaacctgaagatacagcctttattattgtgcc 311
DB 241 ACGGTGTATCTCAAAATGAACAGCCTTGAACCTTGAAGATACGGCGCTTTATTATTGTGCA 300
QY 312 gctgcagcgtccgctggatgatatttccctgcgcggttgggtttgactactggggccag 371
DB 301 GTTCGGCCGTCGGCTGGAGCAGACATTTCCACCCGGATGGTGTGTTTACTGGGGCCAG 360
QY 372 gggaccaggtcacctctcctca 395
DB 361 GGGACCCAGGTACCGCTCTCCTCA 384

RESULT 8
LGL236101
LOCUS LGL236101 384 bp DNA linear MAM 10-FEB-1999
DEFINITION Lama glama mRNA for Ig heavy chain variable region (r3).
ACCESSION AJ236101
VERSION AJ236101.1 GI:4165533
KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable region.
SOURCE llama.
ORGANISM Lama glama
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
REFERENCE 1 (bases 1 to 384)
AUTHORS Frenken, L.G.J.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS

FEATURES
source
1..384
/organism="Lama glama"
/db_xref="taxon:9844"
/cell_type="peripheral blood lymphocyte"
/rearranged
V_region
1..384
/gene="r3"
gene
1..384
/gene="r3"
CDS
1..384
/gene="r3"
/codon_start=1
/protein_id="CAAL5413.1"

/db_xref="GI:4165534"
/translation="QVQLQSGGGLVQAGSLRLSCVASGRTSHCYGCGWFRQVP
GKERELVAARIWMSGLETHKDSVGRFTISRDNKSTVYLQNSLKPEDTAVYICAAAR
TVRVDDISPPGVFGQGTQVTVSS"
BASE COUNT 82 a 97 c 121 g 84 t
ORIGIN

Query Match 57.6%; Score 321.6; DB 4; Length 384;
Best Local Similarity 89.8%; Pred. No. 1.2e-90;
Matches 345; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 12 cagggtgcagctcagagagtcaggggagagattggtgcagctcgggggctctctgagactc 71
DB 1 CAGGTGCAGCTCAGCAGTCAGGGGGAGGATTGGTGCAGGCTGGGGGACTCTCTGAGACTC 60
QY 72 tctgtgcagctcgggacgcgccaccagtggtcagctcagctcagctcagctcagctc 131
DB 61 TCTGTGTAGCTTCGGGACGCGACCTTCAGTGGCCATGGTGTGCGCATGGCTGGCTGTTTC 120
QY 132 cgccaggttccagggaagagcgtgagttgtcgcagctcattaggtggagtggttaagag 191
DB 121 CGCCAGGTTCCAGGGAAGCAGCGTGAGTTGTGCGCAGCTATTAGGTGGAGCGGTCTTGAA 180
QY 192 acatggtataaagactccgtgaaggccagattccatctccagagataaacgccaagact 251
DB 181 ACATATTATGAAGACTCCCGTGAAGGGCCGATTCCACCATCTCCAGAGACAAACCAAGAAC 240
QY 252 acggtttatctcaaatgaacagcctgaacctgaagatacagcctttattattgtgcc 311
DB 241 ATGGTGTATCTCAAAATGAACAGCCTTGAACCTTGAAGATACGGCGCTTTATTATTGTGCA 300
QY 312 gctgcagcgtccgctggatgatatttccctgcgcggttgggtttgactactggggccag 371
DB 301 GCTCGGACGGTTCGGCTGGTGTGACATTTCCAGCCCGTGGTGGTTCCTACTGGGGCCAG 360
QY 372 gggaccaggtcacctctcctca 395
DB 361 GGGACCCAGGTACCGCTCTCCTCA 384

RESULT 9
LGL236099
LOCUS LGL236099 384 bp DNA linear MAM 10-FEB-1999
DEFINITION Lama glama mRNA for Ig heavy chain variable region (r1).
ACCESSION AJ236099
VERSION AJ236099.1 GI:4165529
KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable region.
SOURCE llama.
ORGANISM Lama glama
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
REFERENCE 1 (bases 1 to 384)
AUTHORS Frenken, L.G.J.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS

FEATURES
source
1..384
/organism="Lama glama"
/db_xref="taxon:9844"
/cell_type="peripheral blood lymphocyte"
/rearranged
V_region
1..384
/gene="r1"
gene
1..384
/gene="r1"
CDS
1..384
/gene="r1"
/codon_start=1
/protein_id="CAAL5411.1"

```

/db_xref="GI:4165530"
/translation="QVQLQSGGGGLVQAGDSLRLSCEASGRTSHGYGGYGMGFQVPP
GKERELVAAIRFSGVETHYKDSVKGRTISRDNKNMVLQMNLSLAPEDTAVYICAR
TVRVDLSSPFGFYAGWGQGVQTVSS"
81 a 98 c 122 g 83 t
BASE COUNT
ORIGIN

```

BASE COUNT	81 a	98 c	122 g	83 t
ORIGIN				

Query Match 57.3%; Score 320; DB 4; Length 384;
Best Local Similarity 89.6%; Pred. No. 3.8e-90;
Matches 344; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 12 caggtgcagctgcaggagtcagggggagagattggtgcaggctggggctctctgagactc 71
|||||
Db 1 CAGGTGCAGTGCAGCAGTCAGGGGGGAGGATTGGTGCAGGCTGGGGACTCTCTGAGACTC 60

Qy 72 tcctgtcagcctcgggacgcgccaccagtggtcattggtcactatggtatgggctgttc 13

Db 61 TCCTGTAAGCCTCGGGACGACCTCCATGGCTATGGTGGCTATGGCATGGGCTGGTTC 120

QY 132 cgccaggttccagggaagagcgtgagtttgcgcagctattaggtggagtggtaaaagag 191

Db 121 CGCCAGGTTCCAGGGAAAGCGTGAGCTTGTCCAGCTATTAGGTGGAGCGGTGTGAA 180

Qy 192 acatggataaagactccgtgaaggccgattcaccatctccagagataacgccaaagact 251

Db 181 ACATATCATAAAGACTCCGTGAAGGGCCGGTTCACCATCTCCAGACAGCAACGCCAAGAAC 240

Qy 252 acggttatctgaaaatgaacagcctgaaacccctgaagatacggccgtttattattgtgcc 311

DB 241 ATGGTGTATCTACAAATGAACAGGCTCAAAACCTGAGGACACGGCCGTTATTATTGTGCA 300

[illegible]

```
0v 372 gggaccaggtcaccggtctctca 395
```

Db 361 GGGACCCAGGTCACCGTCTCCTCA 384

RESULT 10
LGL236103

RESULT	10
LGL236103	
LOCUS	LGL236103
DEFINITION	Lama glama mRNA for Ig heavy chain variable region (r5).
ACCESSION	AJ236103
VERSION	AJ236103.1
KEYWORDS	GI:4165537 immunoglobulin heavy chain; immunoglobulin superfamily; variable

SOURCE
Region.
Llama.
ORGANISM
Lama glama
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
REFERENCE
1 (bases 1 to 384)
AUTHORS
Frenken, L.G.J.
TITLE
Direct Submission
JOURNAL
Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever
Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE
NETHERLANDS

V_region	gene	CDS
1..>384	1..384	1..>384
/gene="r5"	/gene="r5"	/gene="r5"

```
/gene="r5"  
/codon_start=1  
/protein_id="CAAL5415.1"
```

```

/db_xref="GI:416538"
/translation="QVQLQSGGGLVQAGGSLRLSCAASGRTSHGYGGYGMGFVQVPP
CKRELVAATRWSGTSTYVADSYKGRFTISRDNKVNVYLQMNLSLQAPDTAVYHCAR
TVRVDISSPVGFGAYGQGQGVTVSS"
77 a 100 c 125 g 82 t
BASE COUNT
ORIGIN

```

BASE COUNT	77 a	100 c	125 g	82 t
ORIGIN	IRVNDICSSFVGRNRGQGIYVSS			

Query Match	57.1%	Score 318.4;	DB 4;	Length 384;
Best Local Similarity	89.3%;	Pred. No. 1.2e-89;		
Matches 343;	Conservative	0;	Mismatches 41;	Indels 0;
				Gaps 0;

Qy 12 caggtgcagctgcaggagtcagggggaggaattggtgcaggtcgtggggctctctgagactc 71
|||||
Db 1 CAGGTGCAGCTGCAGGAGTCAGGGGGAGGATTGTCAGGCTGGGGGCTCTCTGAGACTC 60

Qy 72 tcctgtcagcctcgggacgcccaccagtggtcattggtcactatggtatgggctggtc 13

Db 61 TCCTGTGCAGCCTCGGGACGCACTTCCCATGGCTATGGTGCTATGGCATGGCTGGTTC 120

[illegible]

Db 121 CGCCAGGTTCCAGGGAAGGCGGTGAGCTTGTCCAGCAATTAGGTGGAGCGGTACTAGT 180

Qy 192 acatggataaagactccgtgaaggccgattccaccatctccagagataacgccaaagact 251

Db 181 ACATACTATGCAGACTCCGTGAAGGCCGATTCCACCATCTCCAGAGACAACGTCAAGAAC 240

QY 252 acggtttatctgcaaatgaacagccctgaaccctgaagatacggccgtttattattgtgcc 311

DB 241 ATGGTGTATCTGCAAAATGACACAGCCTTGAATACCTGAGGACACAGGGCCGTTTACCATTGTGCA 300

[illegible]

372 aaaccacaggtcacctctctctca 395

Db 361 GGGACCCAGGTCACCGTCTCCTCA 384

RESULT 11
LGL236104

RESULT	11
LGL236104	
LOCUS	
DEFINITION	Lama glama mRNA for Ig heavy chain variable region (r6).
ACCESSION	AJ236104
VERSION	AJ236104.1 GI:4165539
KEYWORDS	immunoglobulin heavy chain; immunoglobulin superfamily; variable

SOURCE	region.
ORGANISM	llama.
	Lama glama
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
REFERENCE	1 (bases 1 to 384)
AUTHORS	Frenken, L.G.J.
TITLE	Direct Submission
JOURNAL	Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 At Vlaardingen, THE NETHERLANDS

v_region	/rearranged <1. .>384 /gene="r6"
gene	.1. .384 /gene="r6"
CDS	<1. .>384

```

/genes="rf"
/codon_start=1
/protein_id="CAAL5416.1"

```

/db_xref="GI:4165540"
/translation="OVQIQSGGGGLVQAGDSLRSLSCAASGRTHSGYGGYGMGWFQVP
GKERLVAARIWSGNTYYADSVKGRFTISRDNVKNMVLQMSLKPEDTAVYTCADR
TVRVVDISSPVGFAYWGQGTQVTVSS"
BASE COUNT 76 a 97 c 125 g 86 t
ORIGIN

Query Match 56.5%; Score 315.2; DB 4; Length 384;
Best Local Similarity 88.8%; Pred. No. 1.3e-88;
Matches 341; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 12 cagggtcagctcagagtcagggtgaggtggtgagcagctgggggctctctgagactc 71
DB 1 CAGGTGCAGCTCAGCAGCTCGGGGAGGATTGGTGCAGACTGGGGGCTCTCTGAGACTC 60
QY 72 tctgtgagctcgggacgcccaccaggtgctcagctatgtagtggctggttc 131
DB 61 TCCTGTGACGCTCGGGAGCGCACTTCCCATGGGTATGGTGGCTATGGCATGGGCTGGTTC 120
QY 132 cgcacaggttccaggaagagcgtgagttgtcgcagctattaggtggagtggtgaaagag 191
DB 121 CGCCAGGTTCCAGGAGGAGCGTGAGCTTGTCGACAGCAATTAGGTGGAGCGGTATTAGT 180
QY 192 acatggtataagactccgtgaagggccgattccaccatctccagagataacgccaagact 251
DB 181 ACATACTATGACAGCTCCGTCGAAGGGCCGATTCCACCATCTCCAGAGACAACGTCGAAGA 240
QY 252 acggtttatctcaaatgaacagcctgaacctgaagatacagccgtttattattgtgcc 311
DB 241 ATGGGTATCTGCAATGACAGCTTGAAACCTTGAGGACACGCGCCGTTTATTATTGTGCA 300
QY 312 gctcagccgtccgctgagatatttccctgcgctggttgggtttgactactggggccag 371
DB 301 GCTCGACGCTCGCGTGTGACATTTCCAGCCCGGTTGGGTTTGCTACTGGGCCAG 360
QY 372 gggaccaggtcacgctctcctca 395
DB 361 GGGACCCAGTCCCGTCTCCTCA 384

RESULT 12
LGL236105 LGL236105 384 bp DNA linear MAM 10-FEB-1999
DEFINITION Lama glama mRNA for Ig heavy chain variable region (r7).
ACCESSION AJ236105
VERSION AJ236105.1 GI:4165541
KEYWORDS immunoglobulin heavy chain; immunoglobulin superfamily; variable region.
SOURCE llama.
ORGANISM Lama glama
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
REFERENCE 1 (bases 1 to 384)
AUTHORS Frenken, L.G.J.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-1998) Frenken L.G.J., Biotechnology, Unilever Research, Olivier van Noortlaan 120, 3133 AT Vlaardingen, THE NETHERLANDS
FEATURES
source location/Qualifiers
1..384
/organism="Lama glama"
/db_xref="taxon:9844"
/cell_type="peripheral blood lymphocyte"
/rearranged
<1..>384
/gene="r7"
1..384
/gene="r7"
<1..>384
/gene="r7"
/codon_start=1
/protein_id="CAA15417.1"

/db_xref="GI:4165542"
/translation="OVQIQSGGGGLVQAGDSLRSLSCAASGRTHSGYGGYGMGWFQVP
GKERLVAARIWSGNTYYADSVKGRFTISRDNVKNMVLQMSLKPEDTAVYTCADR
TVRVVDISSPVGFAYWGQGTQVTVSS"
BASE COUNT 78 a 96 c 124 g 86 t
ORIGIN

Query Match 53.6%; Score 299.2; DB 4; Length 384;
Best Local Similarity 86.2%; Pred. No. 1.5e-83;
Matches 331; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 12 cagggtcagctcagagtcagggtgaggtggtgagcagctgggggctctctgagactc 71
DB 1 CAGGTGCAGCTCAGCAGCTCAGGGGAGGATTGGTGCAGCTGGGGACTCTCTGAGACTC 60
QY 72 tctgtgagctcgggacgcccaccaggtgctcagctatgtagtggctggttc 131
DB 61 TCCTGTGACGCTCGGGAGCGCACTTCTCATGGGTATGGTGGCTATGGCATGGGCTGGTTC 120
QY 132 cgcacaggttccaggaagagcgtgagttgtcgcagctattaggtggagtggtgaaagag 191
DB 121 CGCCAAATTCAGGAGGAGCGTGAGCTTGTCGACAGCAATTAGGTGGAGCGGTCTGTAAT 180
QY 192 acatggtataagactccgtgaagggccgattccaccatctccagagataacgccaagact 251
DB 181 ACATACTATGACAGCTCCGTCGAAGGGCCGATTCCACCATCTCCAGAGACAACGTCGAAGA 240
QY 252 acggtttatctcaaatgaacagcctgaacctgaagatacagccgtttattattgtgcc 311
DB 241 ATGGGTATCTGCAATGACAGCTTGAAACCTTGAGGACACGCGCCGTTTACACTTTGTGCA 300
QY 312 gctcagccgtccgctgagatatttccctgcgctggttgggtttgactactggggccag 371
DB 301 GTCGACGCTCGCGTGTGACATTTCCAGTCCGCTGGGTTTGCTACTGGGCCAG 360
QY 372 gggaccaggtcacgctctcctca 395
DB 361 GGGACCCAGTCCCGTCTCCTCA 384

RESULT 13
AX021690 AX021690 384 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 21 from Patent WO9923221.
ACCESSION AX021690
VERSION AX021690.1 GI:10044968
KEYWORDS llama.
SOURCE Lama glama
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
REFERENCE 1 (bases 1 to 384)
AUTHORS Howell, S., Ledebor, A.M., Frenken, L.G. and van der Logt, C.P.E.
TITLE Multivalent antigen-binding proteins
JOURNAL Patent: WO 9923221-A 21 14-MAY-1999;
HOWELL STEVEN (GB); UNILEVER PLC (GB); LEDEBOER ADRIANUS MARINUS (NL); LOGT CORNELIS PAUL ERIK V D (NL); UNILEVER NV (NL); FRENKEN LEON GERARDUS JOSEPH (NL)
FEATURES
source location/Qualifiers
1..384
/organism="Lama glama"
/db_xref="taxon:9844"
<1..>384
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07784.1"
/db_xref="GI:10044968"
/translation="OVQIQSGGGGLVQAGDSLRSLSCAASGRTHSGYGGYGMGWFQVP
GKERLVAARIWSGNTYYADSVKGRFTISRDNVKNMVLQMSLKPEDTAVYTCADR
TVRVVDISSPVGFAYWGQGTQVTVSS"
BASE COUNT 78 a 96 c 124 g 86 t
ORIGIN

Qy	131	ccgccaggttccagggaagcgtgagttgttcgcagctattaggtgagtggttaaga	190			
Db	459	CCGCAAAATTTCCAGGGAAGGAGCGGTGAGCTTGTCCAGCAATTAGGTGGAGCGGTCGTAA	518			
Qy	191	gacatgtataaagactccgtgaaggcccatccacatccagagataacgcccaagac	250			
Db	519	TACATACATATCGACACATCCGTAAGGCGCGGATTCACCATCTCCAGAGACAACGTCGAAGGA	578			
Qy	251	tacggtttatctgcaaatgaacagcctgaaacctgaagatacagccgctttattattgtgc	310			
Db	579	CATCCTGTATCTGCAATGAACAGTTTTGAAACCTGAGGACACGGCCGTTTACACTTGTGC	638			
Qy	311	cgctgacccggtccgcgtgcatgatatttccctcgcgcggttgggtttgactactcagggccca	370			
Db	639	AGTTTCGACCGTCCGCGCTGTGTCATTTCCAGTCCCGTTGGGTTTGCCTACTCGGGCCA	698			
Qy	371	gggacccaggtcacc	386			
Db	599	GGGACCCAGGTCACC	714			
RESULT 15						
AX021721						
LOCUS	AX021721	717 bp	DNA linear PAT 07-SEP-2000			
DEFINITION	Sequence 52 from Patent WO923221.					
ACCESSION	AX021721					
VERSION	AX021721.1	GI:10044988				
KEYWORDS	synthetic construct.					
SOURCE	synthetic construct.					
ORGANISM	artificial sequence.					
REFERENCE	1 (bases 1 to 717)					
AUTHORS	Howell,S., Ledebouer,A.M., Frenken,L.G. and van der Logt,C.P.E.					
TITLE	Multivalent antigen-binding proteins					
JOURNAL	Patent: WO 923221-A 52 14-MAY-1999;					
	HOWELL, STEVEN (GB); UNILEVER PLC (GB); LEDEBOER ADRIANUS MARINUS					
	(NL); LOGT CORNELIS PAUL ERIK V D (NL); UNILEVER NV (NL); FRENKEN					
	LEON GERARDUS JOSEPH (NL)					
FEATURES	Location/Qualifiers					
source	1..717					
	/organism="synthetic construct"					
	/db_xref="taxon:32630"					
	/note="plasmid construct; double stranded; linear"					
BASE COUNT	158 a	183 c	220 g 156 t			
ORIGIN						
Query Match 52.2%; Score 291.2; DB 6; Length 717;						
Best Local Similarity 85.9%; Pred. No. 5.4e-81;						
Matches 323; Conservative 0; Mismatches 53; Indels 0; Gaps 0;						
Qy	11	acaggtcagctgcagagtcaggggagagattggttcagcgtgggggctctctgagact	70			
Db	342	ACAGGTCCAGTGCAGCAGTCAGGGGCGAGGATTGGTCAGGCTGGGGACTCTCTGAGACT	401			
Qy	71	ctcctgtcagctcgggagcgcgccaccagtggtcgtcactatgtagtggctggtt	130			
Db	402	CTCCTGCGCGGCTTCGGGAGCGCACTTCTCATGGGTATGGGTATGGCATGGGCTGGTT	461			
Qy	131	ccgcaggttccaggggaagcagcgtgagttgttcgcagctattaggtggagtgtaaga	190			
Db	462	CCGCAAAATTTCCAGGGAAGGAGCGGTGAGCTTGTCCAGACAATTAGGTGGAGCGGTCGTAA	521			
Qy	191	gacatgtataaagactccgtgaaggcccatccacatccagagataacgcccaagac	250			
Db	522	TACATACATATCGACACTCCGTAAGGCGCGGATTCACCATCTCCAGAGACAACGTCGAAGGA	581			
Qy	251	tacggtttatctgcaaatgaacagcctgaaacctgaagatacagccgctttattattgtgc	310			
Db	582	CATGCTCTATCTGCAATGAACAGTTTGAACCTGAGGACACGGCCGTTTACACTTGTGC	641			
Qy	311	cgcctgacccggtccgcgtgcatgatatttccctcgcgcggttgggtttgactactcagggccca	370			

Db 642 AGTTCGGACGGTCCGCGTGTGACATTTCCAGTCCGGTTGGGTTTGCCCTACTGGGGCCA 701

Qy 371 ggggaccaggtcacc 386

|||||

Db 702 GGGGACCCAGGTCACC 717

Search completed: August 13, 2002, 12:30:12
Job time: 20575 sec

OM of: US-09-742-690-1 to: Issued_Patents_AA.* out_format : pfs

Date: Aug 13, 2002 1:42 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=frame+np2.model -DRV=xlp
-Q/cgn2_1/USPTO.spool/US09742690/runat_12082002_094914_87/app_query.fasta_1.623
-DB=Issued_Patents_AA -OFMT=fastan -SUFFIX=ra1 -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=us09742690.ecn1_18 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1
```

Search information block:

```
Query: US-09-742-690-1
Query length: 558
Database: Issued_Patents_AA.*
Database sequences: 231828
Database length: 24425594
Search time (sec): 39.700000
```

score_list:

```
Sequence Strd Orig ZScore EScore Len ! Documentation
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-983-607-28 + 403.50 706.85 131 !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-974-899-6 + 399.50 701.29 113 ! S
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-079-029-10 + 395.00 682.93 312 !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-129-9308-96 + 394.00 689.45 4.6e-31 139 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-134-346A-51 + 394.00 689.45 4.6e-31 139 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-976-288A-96 + 394.00 689.45 4.6e-31 139 !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-07-988-925-11 + 393.50 690.15 4.9e-31 119 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-362-780-11 + 393.50 690.15 4.9e-31 119 !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-07-934-373B-21 + 393.00 689.01 5.6e-31 122 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-437-642B-21 + 393.00 689.01 5.6e-31 122 !
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US93-07832-21 + 393.00 689.01 5.6e-31 122 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-025-769B-178 + 392.50 679.58 8.1e-31 281 !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-331-398A-46 + 391.50 686.62 7.7e-31 119 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-331-397B-46 + 391.50 686.62 7.7e-31 119 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-759-804A-46 + 391.50 686.62 7.7e-31 119 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-227-693-46 + 391.50 686.62 7.7e-31 119 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-552-816A-14 + 391.00 685.99 8.6e-31 116 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-862-124-17 + 390.50 675.83 1.3e-30 287 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-862-124-14 + 390.50 675.24 1.3e-30 304 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-958-201-2 + 390.00 684.05 1.1e-30 118 ! S
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-958-201-4 + 390.00 684.05 1.1e-30 118 ! S
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US96-13152-4 + 389.00 668.72 2.1e-30 443 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-025-769B-38 + 388.00 680.34 1.7e-30 120 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-025-769B-63 + 388.00 680.34 1.7e-30 120 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-983-607-32 + 387.50 677.88 2.0e-30 140 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-983-607-36 + 387.00 678.92 2.1e-30 116 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-976-183A-55 + 386.50 675.40 2.6e-30 150 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-665-202-30 + 385.50 675.67 3.1e-30 123 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-428-197-1 + 385.50 675.50 3.1e-30 125 ! S
/cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US93-10555-1 + 385.50 675.50 3.1e-30 125 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-579-378A-20 + 385.00 673.83 3.5e-30 135 !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-398-612A-30 + 385.00 667.47 4.3e-30 251 !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-398-611A-30 + 385.00 667.47 4.3e-30 251 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-931-334A-30 + 385.00 667.47 4.3e-30 251 !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-09-027-449-27 + 385.00 667.47 4.3e-30 251 !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-804-444A-27 + 385.00 667.47 4.3e-30 251 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-026-985-27 + 385.00 667.47 4.3e-30 251 !
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-398-613A-30 + 385.00 667.43 4.3e-30 252 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-653-402B-10 + 384.50 671.33 4.1e-30 158 !
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-653-402B-6 + 384.50 671.33 4.1e-30 158 !
```

```
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-983-607-46 + 383.50 672.64 4.7e-30 117
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-07-934-373C-4 + 383.00 671.50 5.3e-30 120
/cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-437-642B-4 + 383.00 671.50 5.3e-30 120
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-752-693A-4 + 382.50 670.87 5.9e-30 117
/cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-652-558-39 + 382.50 670.61 6.0e-30 120
```

seq_name: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-983-607-28

```
seq_documentation_block:
; Sequence 28, Application US/08983607
; Patent No. 6140470
; GENERAL INFORMATION:
; APPLICANT: Xiaohong Cai
; APPLICANT: Alan Garen
; TITLE OF INVENTION: Human Anti-Tumor Monoclonal Anti-
; TITLE OF INVENTION: bodies
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Department of Molecular Biophysics
; ADDRESSEE: and Biochemistry, Yale University
; STREET: 266 Whitney Avenue
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: United States of America
; ZIP: 06520-8114
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" 1.44 Mb diskette
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Word Processing
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/983,607
; FILING DATE: April 27, 1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IB96/01032
; FILING DATE: June 28, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mary M. Krinsky
; REGISTRATION NUMBER: 32423
; REFERENCE/DOCKET NUMBER: OCR-679
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 203-773-9544
; TELEFAX: 203-773-1183
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 131 residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: polypeptide
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens (melanoma patient immu-
; ORGANISM: nized with autologous tumor cells)
; INDIVIDUAL ISOLATE: peripheral blood lymphocytes
; IMMEDIATE SOURCE:
; LIBRARY: DM414 scfv antibodies obtained from
; LIBRARY: fUSE5 fusion phage construct
; CLONE: V13
; FEATURE:
; NAME/KEY: heavy chain
; US-08-983-607-28

alignment_scores:
Quality: 403.50 Length: 161
Ratio: 3.736 Gaps: 4
Percent Similarity: 67.081 Percent Identity: 55.901

alignment_block:
US-09-742-690-1 x US-08-983-607-28
```


Align seg 1/1 to: US-08-983-607-28 from: 1 to: 131

```

12 CAGGTGACGTGACGAGTCCAGGGGAGGAGTGTGTCAGGCTGGGGCTC 61
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
1 GluValGlnLeuValGlnSerGlyGlyGlyLeuValGlnProGlyGlyse 17
62 TCTGAGACTCTCTGTGTCAGCTCGGACCGGACCCAGCCAGCTGGTCATGTC 111
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
112 ACTATGCTATGGCTGCTCCGAGCTCCAGGGAGGAGGAGGAGGAGTGT 161
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
162 GTCGACGATTTAGTGGTGGTGTAAAGACATGTTATTAAGACTCCGT 211
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
48 ValSerAlaIleSerGlySerGlyGlySerThrTyrTyrAlaAspSerVa 64
212 GAAGGGCGGATTCACCATCTCCAGATACGCGGACGAGTGTATTAAGTGC 311
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
64 LysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
262 TGCAATGAACAGCTGAAACCTGAAGATACGCGGACGAGTGTATTAAGTGC 311
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
312 .....GCTCGACCGGTCGCGGTGGATGATATTTCCTCGCGGTGGGTT 355
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
98 LysGlyValAlaPro.....Ph 103
356 TGACTACTGGGGCAGGGGACCCAGGTCACCGCTCTCTCAGGATCTCATC 405
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
103 eAspTyrTrpGlyGlnGlyThrProValThrValSerSerGlyGly... 118
406 ACCATCACCATCAGGATCCACCTCCATTTGAAGTCTGTCACCTCCTCAC 455
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
118 ..... 118

```

seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:US-08-974-899-6

```

seq_documentation_block:
; Sequence 6, Application US/08974899
; Patent No. 6037454
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: Humanized Anti-CD11a Antibodies
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,899
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/031971
; FILING DATE: 11/27/96
; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40.378
; REFERENCE/DOCKET NUMBER: P1014R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-974-899-6

```

alignment_scores:

Quality:	399.50	Length:	128
Ratio:	4.035	Gaps:	2
Percent Similarity:	77.344	Percent Identity:	64.844

alignment_block:

```

US-09-742-690-1 x US-08-974-899-6 ..
Align seg 1/1 to: US-08-974-899-6 from: 1 to: 113
12 CAGGTGACGTGACGAGTCCAGGGGAGGAGTGTGTCAGGCTGGGGCTC 61
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
1 GluValGlnLeuValGlnSerGlyGlyGlyLeuValGlnProGlyGlyse 17
62 TCTGAGACTCTCTGTGTCAGCTCGGACCGGACCCAGCCAGCTGGTCATGTC 111
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
112 ACTATGCTATGGCTGCTCCGAGTGTCCAGGAGGAGGAGGAGGAGTGT 161
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
162 GTCGACGATTTAGTGGTGGTGTAAAGACATGTTATTAAGACTCCGT 211
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
48 ValSerValIleSerGlyAspGlyGlySerThrTyrTyrAlaAspSerVa 64
212 GAAGGGCGGATTCACCATCTCCAGATACGCGGACGAGTGTATTAAGTGC 261
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
64 LysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
262 TGCAATGAACAGCTGAAACCTGAAGATACGCGGACGAGTGTATTAAGTGC 311
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
312 GCTCGACCGGTCGCGGTGGATGATATTTCCTCGCGGTGGTGGTACTA 361
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
98 Arg.....GlyPheAspTy 102
362 CTGGGGCAGGGGACCCAGGTCACCGTCTCTCTCA 395
   :::::::::::::::::::: :::::::::::::::::::: ::::::::::
102 rTrpGlyGlnGlyThrLeuValThrValSerSer 113

```

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-09-079-029-10

```

seq_documentation_block:
; Sequence 10, Application US/09079029
; Patent No. 6342369
; GENERAL INFORMATION:
; APPLICANT: Adams, Camilia W.
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Chuntharapal, Anan
; APPLICANT: Kim, Kyung J.
; TITLE OF INVENTION: Apo-2 Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California

```

```

; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genetech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,029
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P1101R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5416
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 312 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-09-079-029-10

alignment_scores:
  Quality: 395.00      Length: 193
  Ratio: 3.211        Gaps: 7
  Percent Similarity: 63.731 Percent Identity: 51.295

alignment_block:
US-09-742-690-1 x US-09-079-029-10 ..
  Align seg 1/1 to: US-09-079-029-10 from: 1 to: 312

15 GTGCAGCTCAGAGTCAGGGGAGATTGGTGCGAGCTGGGGCTCTCT 64
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
41 ValGlnLeuValGlnSerGlyGlyLeuValGlnProGlyGlySerLe 57
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
65 GAGACTCTCTGTGAGCTCGGAGCGGCCACCGCTGCTGCTGCTACT 114
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
57 uArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer.....T 71

115 ATGGTATGGCTGGTTCGCCAGGTCCAGGAAGGAGCGGTGTTGTC 164
|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
71 yTrpMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrpVal 87

165 GCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAAGACTCCGTGAA 214
|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
88 AlaAsnIleLysGlnAspGlySerGluLysTyTrpValAspSerVally 104

215 GGGCCGATTCCACATCTCAGAGATAACGCCAGACTACGGTTATCTGC 264
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
104 sGlyArgPheThrIleSerArgAspAsnAlaLysAsnSerLeuTyLeuG 121

265 AAATGAACAGCTGAACCTCAGAGATACGCCGCTTTATTATTGTCGCGCT 314
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
121 InMetAsnSerLeuArgAlaGluAspThrAlaValTyTrpCysAlaArg 137

315 CGACCGGTCGGGTGATGATATTTCCCTCGCGGTTGGGTTTGACTACTG 364
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
138 AspLeuLysValLysGlySerSer...SerGlyTrp.PheAspProTr 153

365 GGGCCAGGGAGCCAGGTCCAGCTCTCCTCAGGATCTCATCACCATCACC 414
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
153 pGlyArgGlyThrThrValThrValSerSerGlyGly..... 165

415 ATCAGCGATCCACCTCCATGAAGTCTGACCCAGTCTCAGTACGCTCAG 464
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
166 .....GlyGly 167

465 TGTGTGGTATTGGTTACTCCGCT..... 488

```

```

|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
168 SerGlyGlyGlySerGlyGlyGlyGlyGlySerSerGluLeuThrGlnAs 184
489 .CCAACCGTCTGTGCTCT.....GGTACCACCTGTGACG. 521
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
184 pProAlaValSerValAlaLeuGlyGlnThrValArgIleThrCysGlnG 201
522 .....GTTCTGAACCTTACTACTCC 542
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
201 LyAspSerLeuArgSerTyTyAla 209

seq_name: /cgn2_6/ptodata/2/laa/5A_COMB.pep:US-08-129-930B-96

seq_documentation_block:
; Sequence 96, Application US/08129930B
; Patent No. 5804187
; GENERAL INFORMATION:
; APPLICANT: do Couto Dr., Fernando J.R.
; APPLICANT: Ceriani Dr., Roberto L.
; APPLICANT: Peterson Dr., Jerry A.
; APPLICANT: Padian Dr., Eduardo A.
; TITLE OF INVENTION: Analogue Peptides with Broad
; TITLE OF INVENTION: Carcinoma Specificity, and Kit and
; TITLE OF INVENTION: Diagnostic Vaccination and
; TITLE OF INVENTION: Therapeutic Methods
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: V. AMZEL & ASSOC.
; STREET: 2055 No. 5804187th Broadway, Suite 201
; CITY: Walnut Creek
; STATE: California
; COUNTRY: USA
; ZIP: 94596
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/129,930B
; FILING DATE: September 30, 1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Amzel Ph.D., Viviana
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: CRFCC-008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 521-1333
; TELEFAX: (510) 521-3541
; TELEX: n.a.
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-129-930B-96

alignment_scores:
  Quality: 394.00      Length: 128
  Ratio: 3.863        Gaps: 2
  Percent Similarity: 79.688 Percent Identity: 63.281

alignment_block:
US-09-742-690-1 x US-08-129-930B-96 ..
  Align seg 1/1 to: US-08-129-930B-96 from: 1 to: 139

12 CAGGTGCAGCTGCAGGAGTCAGGGGAGGATTGGTCAGGCTGGGGCTC 61
:::|||||:::  |||||  |||||  |||||  |||||  |||||  |||||  |||||
20 GluValGlnMetValGluSerGlyGlyGlyLeuValGlnProGlyGlySe 36

```

```

alignment_scores:
  Quality: 394.00      Length: 128
  Ratio: 3.863        Gaps: 2
  Percent Similarity: 79.688      Percent Identity: 63.281

alignment_block:
US-09-742-690-1 x US-08-134-346A-51      ..

Align seg 1/1 to: US-08-134-346A-51 from: 1 to: 139

12  CAGGTCGACGTCGAGGAGTCAGGGGGAGGATTGGTCAGGCTGGGGGCTC 61
    ::::::::::: |||||
20  GluValGlnMetValGluSerGlyGlyGlyLeuValGlnProGlyGlySer 36
    |||||
62  TCTGAGACGTCCTCTGTGCAGCCTCTGGGACGGCCACCACTGTCATGGTC 111
    |||||
36  rLeuArgLeuSerCysAlaAlaSerGlyPheAlaPheSerSer..... 50

112 ACTATGGTATGGGCTGGTCCGCCAGGTTCAGGAGGAGCGTCAGTTT 161
    |||||
51  ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
    |||||
162 GTCGCGACGCTATTAGGTGGAGTGTAAAGAGACATGGTATAAGACTCCGT 211
    |||||
67  ValAlaGluIleSerSerGlyGlyAsnTyrAlaTyrTyrGlnAspThrVa 83

212 GAAGGGCCGATTCCACCATCCAGAGATAAGCCCAAGACTACGGTTTATC 261
    |||||
83  lThrGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 100

262 TGCAAATGAACAGCCTGAAACCTTGAAGATACGGCCGTTTATTATTGTGCC 311
    |||||
100 euGlnMetAsnSerLeuArgAlaGluaspThrAlaValTyrTyrCysAla 116

312 GCTCGACCGGTCGCGGTGGATGATATTTCCTCGCGGTTGGGTTTGACTA 361
    |||||
117 Arg.....GluaspTyrGlyIleProAlaTrpPheAlaTyr 128

362 CTGGGCCGACGGGACCCAGGTCCACCGTCTCCCTCA 395
    |||||
128 rTrpGlyGlnGlyThrLeuValThrValSerSer 139

seq_name: /cqn2_6/ptodata/2/iaa/6B_COMB_pep:US-08-976-288A-96

```

```

362 CTTGGGGCCAGGGGACCCAGGTCCACCGTCTCCTCA 395
      |||||
128 rtrpGlyGinglyThrLeuValThrValSer 139

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-976-288A-96

seq_documentation_block:
; Sequence 96, Application US/08976288A
; Patent No. 6315997
; GENERAL INFORMATION:
; APPLICANT: do Couto Dr., Fernando J.R.
; APPLICANT: Ceriani Dr., Roberto L.
; APPLICANT: Peterson Dr., Jerry A.
; APPLICANT: Padlan Dr., Eduardo A.
; TITLE OF INVENTION: Analogue Peptides With Broad
; TITLE OF INVENTION: Carcinoma Specificity, and Kit and
; TITLE OF INVENTION: Diagnostic Vaccination and
; TITLE OF INVENTION: Therapeutic Methods
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder & Poplawski
; STREET: 444 South Flower St., 19th Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM: ..
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM-PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS 5.0
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,288A
; FILING DATE: No. 6315997ember 21, 1997

```

```

; CLASSIFICATION: 424
; PRIOR APPLICATION DATA: 08/129,930
; FILING DATE: September 30, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,696
; FILING DATE: No. 6315997ember 16, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel Ph.D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: P6639938
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 622-7700
; TELEFAX: (213) 489-4210
; TELEX: n.a.
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 139 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-976-288A-96

alignment_scores:
  Quality: 394.00      Length: 128
  Ratio: 3.863        Gaps: 2
  Percent Similarity: 79.688      Percent Identity: 63.281

alignment_block:
US-09-742-690-1 x US-08-976-288A-96 ..

Align seg 1/1 to: US-08-976-288A-96 from: 1 to: 139

12 CAGGTGCGAGTGCAGGAGTGCAGGGGAGGAGTGGTCAGCGCTGGGGCTC 61
  ::::::::::::::::::::|
20 GluValGlnMetValGluSerGlyGlyGlyValGlnProGlyGlyse 36
  ::::::::::::::::::::|
62 TCRGAGACTCTCTGTGCAGCTCGGGACCGCCAGCATGGTGCATGGTC 111
  ::::::::::::::::::::|
36 rLeuArgLeuSerCysAlaAlaSerGlyPheAlaPheSerSer..... 50
  ::::::::::::::::::::|
112 ACTATGTTATGGCTGTTCCGCGAGTTCAGGAGGAGGAGCGTGGT 161
  ::::::::::::::::::::|
51 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 66
  ::::::::::::::::::::|
162 GTGCGACCTATTAGTGGTGGTGTAAAGAGACATGGTATAAGACTCCGT 211
  ::::::::::::::::::::|
67 ValAlaGluSerSerGlyGlyAsnTyrAlaTyrTrpGlnAspThrVa 83
  ::::::::::::::::::::|
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGTTTATC 261
  ::::::::::::::::::::|
83 lThrGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 100
  ::::::::::::::::::::|
262 TGCATATGAACAGCTGAACCTGAAGATACGCGCGTTTATTATTGTC 311
  ::::::::::::::::::::|
100 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 116
  ::::::::::::::::::::|
312 GCTCGACCGGTCCGCGGAGTATATTTCCTGCGGTGGTGGTTGACTA 361
  ::::::::::::::::::::|
117 Arg.....GluAspTyrGlyIleProAlaTrpPheAlaTyr 128
  ::::::::::::::::::::|
362 CTGGGGCCAGGGACCGAGTCCACCGTCTCCTCA 395
  ::::::::::::::::::::|
128 rTrpGlyGlnGlyThrLeuValThrValSerSer 139
  ::::::::::::::::::::|

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-07-988-925-11
seq_documentation_block:
; Sequence 11, Application US/07988925
; Patent No 5585097
; GENERAL INFORMATION:
; APPLICANT: Bolt, Sarah L

```

```

; APPLICANT: Clark, Michael R
; APPLICANT: Gorman, Scott D
; APPLICANT: Applledge, Edward G
; APPLICANT: Waldmann, Herman
; TITLE OF INVENTION: antibody preparation
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon and vanderhye pc
; STREET: 11th Floor, 1100 No. 5585097th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/988,925
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9206422.9
; FILING DATE: 24-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB92/01933
; FILING DATE: 21-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C
; REGISTRATION NUMBER: 29009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-988-925-11

alignment_scores:
  Quality: 393.50      Length: 128
  Ratio: 3.935        Gaps: 2
  Percent Similarity: 78.125      Percent Identity: 64.844

alignment_block:
US-09-742-690-1 x US-07-988-925-11 ..

Align seg 1/1 to: US-07-988-925-11 from: 1 to: 119

12 CAGGTGCGAGTGCAGGAGTGCAGGGGAGGAGTGGTCAGGCTGGGGCTC 61
  ::::::::::::::::::::|
1 GluValGlnLeuLeuGluSerGlyGlyGlyValGlnProGlyGlyse 17
  ::::::::::::::::::::|
62 TCRGAGACTCTCTGTGCAGCTCGGGACCGCCAGCATGGTGCATGGTC 111
  ::::::::::::::::::::|
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
  ::::::::::::::::::::|
112 ACTATGTTATGGCTGTTCCGCGAGTTCAGGAGGAGGAGCGTGGT 161
  ::::::::::::::::::::|
32 ..PheProMetAlaTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
  ::::::::::::::::::::|
162 GTCCGACCTATTAGTGGTGGTAAAGAGACATGGTATAAGACTCCGT 211
  ::::::::::::::::::::|
48 ValSerThrIleSerThrSerGlyArgThrTyrTyrArgAspSerVa 64
  ::::::::::::::::::::|
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGTTTATC 261
  ::::::::::::::::::::|
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
  ::::::::::::::::::::|

```

.Align seg 1/1 to: US-08-362-780-11 from: 1 to: 119

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-07-934-373C-21

alignment_scores:
  Quality: 393.00      Length: 128
  Ratio: 3.891        Gaps: 2
Percent Similarity: 78.906 Percent Identity: 65.625

alignment_block:
US-09-742-690-1 x US-07-934-373C-21 ..
Align seg 1/1 to: US-07-934-373C-21 from: 1 to: 122

12 CAGGTGCAGCTGAGGAGTCCAGGGGAGGAGTGTGTCAGGCTGGGGCTC 61
  ::::::::::::::::::::
1 GluValGlnLeuValGluSerGlyGlyLeuValGlnProGlyGlySe 17
  ::::::::::::::::::::
62 TCTGAGACTCTCTGTCAGCTCCAGCTCGGAGCGGCCACAGCTGTCATGTC 111
  ::::::::::::::::::::
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
  ::::::::::::::::::::
112 ACTATGTATGGCTGTGTCAGGAGTCCAGGAGGAGGAGGAGGAGGAGTTC 161
  ::::::::::::::::::::
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
  ::::::::::::::::::::
162 GTCCGACCTATTAGTGGAGTGTGTAAGAGACATGTTATTAAGACTCCGT 211
  ::::::::::::::::::::
48 ValSerValIleSerGlyAspGlyGlySerThrTyrAlaAspSerVa 64
  ::::::::::::::::::::
212 GAAGGCGCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
  ::::::::::::::::::::
64 LysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
  ::::::::::::::::::::
262 TGCAAAATGAACAGCCTGAACCTGAAGATAGCGCGCTTATTATTGTCGC 311
  ::::::::::::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
  ::::::::::::::::::::
312 GCTGACCGGTCCGCGTGATATTTCCCTGCGGTTGGGTTTGACTA 361
  ::::::::::::::::::::
98 ArgGlyArgVal.....GlyTyrSerLeuSerGlyLeuTyrAspTy 111
  ::::::::::::::::::::
362 CTGGGGCCAGGGAGCCAGGTACCGTCTCTCA 395
  ::::::::::::::::::::
111 rTrpGlyGlnGlyThrLeuValThrValSerSer 122
  ::::::::::::::::::::

seq_name: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-437-642B-21

seq_documentation_block:
; Sequence 21, Application US/08437642B
; Patent No. 6054297
; GENERAL INFORMATION:
; APPLICANT: Paul J. Carter
; APPLICANT: Leonard G. Presta
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/437,642B

```

```

; FILING DATE: 09-May-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/146206
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/15272
; FILING DATE: 14-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P0709P2C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1994
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-437-642B-21

alignment_scores:
  Quality: 393.00      Length: 128
  Ratio: 3.891        Gaps: 2
Percent Similarity: 78.906 Percent Identity: 65.625

alignment_block:
US-09-742-690-1 x US-08-437-642B-21 ..
Align seg 1/1 to: US-08-437-642B-21 from: 1 to: 122

12 CAGGTGCAGCTGAGGAGTCCAGGGGAGGAGTGTGTCAGGCTGGGGCTC 61
  ::::::::::::::::::::
1 GluValGlnLeuValGluSerGlyGlyLeuValGlnProGlyGlySe 17
  ::::::::::::::::::::
62 TCTGAGACTCTCTGTCAGCTCCAGCTCGGAGCGGCCACAGCTGTCATGTC 111
  ::::::::::::::::::::
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
  ::::::::::::::::::::
112 ACTATGTATGGCTGTGTCAGGAGTTCAGGAGGAGGAGGAGGAGGAGTTC 161
  ::::::::::::::::::::
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
  ::::::::::::::::::::
162 GTCGACCTATTAGTGGAGTGTGTAAGAGACATGTTATTAAGACTCCGT 211
  ::::::::::::::::::::
48 ValSerValIleSerGlyAspGlyGlySerThrTyrTyrAlaAspSerVa 64
  ::::::::::::::::::::
212 GAAGGCGCGATTACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
  ::::::::::::::::::::
64 LysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
  ::::::::::::::::::::
262 TGCAAAATGAACAGCCTGAACCTGAAGATAGCGCGCTTATTATTGTCGC 311
  ::::::::::::::::::::
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
  ::::::::::::::::::::
312 GCTGACCGGTCCGCGTGATATTTCCCTGCGGTTGGGTTTGACTA 361
  ::::::::::::::::::::
98 ArgGlyArgVal.....GlyTyrSerLeuSerGlyLeuTyrAspTy 111
  ::::::::::::::::::::
362 CTGGGGCCAGGGAGCCAGGTACCGTCTCTCA 395
  ::::::::::::::::::::
111 rTrpGlyGlnGlyThrLeuValThrValSerSer 122
  ::::::::::::::::::::

seq_name: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US93-07832-21

```

```

seq_documentation_block:
; Sequence 21, Application PC/TUS9307832
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: Immunoglobulin Variants
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07832
; FILING DATE: 19930820
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/715272
; FILING DATE: 14-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05126
; FILING DATE: 15-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934373
; FILING DATE: 21-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME:
; REGISTRATION NUMBER:
; REFERENCE/DOCKET NUMBER: 709P2PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; PCT-US93-07832-21

alignment_scores:
Quality: 393.00 Length: 128
Ratio: 3.891 Gaps: 2
Percent Similarity: 78.906 Percent Identity: 65.625

alignment_block:
US-09-742-690-1 x PCT-US93-07832-21 ..
Align seg 1/1 to: PCT-US93-07832-21 from: 1 to: 122

12 CAGTGCAGCTGCAGGAGTCAGGGGAGGATGCTGCAGGCTGGGGCTC 61
:::|||||
1 GiuValGlnLeuValGluSerGlyGlyGlyLeuValGlnProGlyGly 17
:::|||||
62 TCTGAGACTCTCTGTGTCAGCTCGGACGCGCCAGCATGTCATGCTC 111
|||||
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
112 ACTATGATGGCTGTTCCGCCAGGTTCAGGAAGGACGCTGAGTTT 161
|||||
32 ..TyrAlaMetSerTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
|||||
162 GTCGCAGCTATTAGTGCAGTGGTAAAGACATGGTATAAAGACTCCGT 211
|||||
48 ValSerValIleSerGlyAspGlyGlySerThrTyrTyrAlaAspSerVa 64
|||||

```

```

212 GAAGGCCGATTCACCATCTCCAGAGATAACGCCAAGACTACGTTTATC 261
|||||
64 lLysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
|||||
262 TGCAAATGAACAGCCTCAAACTCAAGATACGCGCGTTTATTATTGCCC 311
|||||
81 euGlnMetAsnSerLeuAlaGluAspThrAlaValTyrTyrCysAla 97
|||||
312 GCTGACCGGTCGCGTGGATGATATTCCTCGCGGTTGGGTTTGACTA 361
|||||
98 ArgGlyArgVal.....GlyTyrSerLeuSerGlyLeuTyrAspTy 111
|||||
362 CTGGGCGCAGGCGACCGAGTCAACCGTCTCTCTCA 395
|||||
111 rTrpGlyGlnGlyThrLeuValThrValSerSer 122
|||||
seq_name: /cqn2_6/ptodata/2/1aa/6B_COMB.pep:US-09-025-769B-178
seq_documentation_block:
; Sequence 178, Application US/09025769B
; Patent No. 6300064
; GENERAL INFORMATION:
; APPLICANT: Knappik, Achim
; APPLICANT: Pack, Peter
; APPLICANT: Ilag, Vic
; APPLICANT: Ge, Liming
; APPLICANT: Moroney, Simon
; APPLICANT: Plueckthun, Andreas
; TITLE OF INVENTION: Protein/(Poly)peptide libraries
; NUMBER OF SEQUENCES: 373
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James F. Haley, Jr., Esq. c/o Fish & Neave
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10021
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/025,769B
; FILING DATE: 18-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95 11 3021.0
; FILING DATE: 18-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: James F. Haley, Jr., Esq.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: MORPHO/5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)596-9000
; TELEFAX: (212)596-9090
; INFORMATION FOR SEQ ID NO: 178:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-025-769B-178

alignment_scores:
Quality: 392.50 Length: 164
Ratio: 3.634 Gaps: 4
Percent Similarity: 65.854 Percent Identity: 55.488

alignment_block:
US-09-742-690-1 x US-09-025-769B-178 ..
Align seg 1/1 to: US-09-025-769B-178 from: 1 to: 281

```



```

; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Benhar, Itai
; TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-
; TITLE OF INVENTION: Specific Antibody Fragments, Fusion Proteins, and Uses
; TITLE OF INVENTION: thereof
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Street Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/331,397B
; FILING DATE: 28-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/767,331
; FILING DATE: 30-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/596,289
; FILING DATE: 12-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 015280-126120US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..119
; OTHER INFORMATION: /note= "Human fetal immunoglobulin
; OTHER INFORMATION: 56Pl'CL Variable Heavy chain (V-H)"
US-08-331-397B-46

alignment_scores:
    Quality: 391.50      Length: 128
    Ratio: 3.876        Gaps: 2
Percent Similarity: 78.906      Percent Identity: 64.062

alignment_block:
US-09-742-690-1 x US-08-331-397B-46

Align seg 1/1 to: US-08-331-397B-46 from: 1 to: 119

12 CAGTGCAGCTGCAGGAGTACAGGGGAGGATGTGTGCAGGCTGGGGCTC 61
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 GlnValGluLeuValGluSerGlyGlyValValGlnProGlyArgSe 17

62 TCTGAGACTCTCCTGTGCAGCTCGGACGGCCGCCACCATGGTATGGTC 111
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 rLeuArgLeuSerCysAlaLaSerGlyPheThrPheSerSer..... 31

112 ACTATGTTGTTGGCTGGTTCCTCCAGGAGGAGGAGGAGGAGGAGTTT 161
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 ..TyrAlaMethIstPValargGlnAlaProGlyLysGlyLeuGluTrp 47
```

```

162 GTCACAGCTATTAGTGGAGTGTAAAGACAGACATGGTATAAGACTCCGT 211
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
48 ValAlaValIleSerTyrAspGlySerAsnLysTyrTyrAlaAspSerVa 64

212 GAAGGGCCGATTCACCATCTCCAGATACAGCCACAGACTACGGTTTATC 261
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
64 LysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81

262 TGAATGAACAGCCTGAAACCTCAAGATACAGCGCGTTTATTATTGTGCC 311
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97

312 GCTGACCCGTCGGCGTGGATGATATTTCCCTGCCCGTGGGTTTGACTA 361
|||||:|||||:|||||:|||||:|||||:|||||:
98 ArgArgSerAlaArgThr.....TyrTyrPheAspTy 108

362 CTGGGGCCAGGGACCCAGGTCACTCCCTCTCTCA 395
|||||:|||||:|||||:|||||:|||||:|||||:
108 rTrpGlyGlnGlyThrLeuValThrValSerSer 119

seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-759-804A-46
seq_documentation_block:
; Sequence 46, Application US/08759804A
; Patent No. 5990296
; GENERAL INFORMATION:
; APPLICANT: Pastan, Ira
; APPLICANT: Willingham, Mark
; APPLICANT: Fitzgerald, David J.
; APPLICANT: Brinkmann, Ulrich
; APPLICANT: Pal, Lee
; TITLE OF INVENTION: Tumor-Specific Antibody Fragments,
; TITLE OF INVENTION: Fusion Proteins, and Uses Thereof
; NUMBER OF SEQUENCES: 68
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/759,804A
; FILING DATE: 03-DEC-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/331,398
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/767,331
; FILING DATE: 30-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/596,289
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen L.
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 015280-126140US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..119
; OTHER INFORMATION: /note="Human fetal immunoglobulin
; OTHER INFORMATION: 56P1'CL Variable Heavy chain (V-H)"
US-08-759-804A-46
```

```
alignment_scores:
  Quality: 391.50      Length: 128
  Ratio: 3.876        Gaps: 2
Percent Similarity: 78.906      Percent Identity: 64.062
```

```
alignment_block:
US-09-742-690-1 x US-08-759-804A-46 ..
```

```
Align seg 1/1 to: US-08-759-804A-46 from: 1 to: 119
```

```
12 CAGTGCAGCTGCAGGAGTCCAGGGGAGGATTGGTCAGGCTGGGGGCTC 61
|||||:||||| |||||||||||||||:||||| ||| ||
1 GlnValGluLeuValGluSerGlyGlyValValGlnProGlyArgse 17
62 TCTGAGACTCTCTGTGCAGCTCGGCGGCGCCAGTCAGTGGTCATGGTC 111
|||||:||||| |||||||||||||||:|||||
17 rLeuArgLeuSerCysAlaAlaSerGlyPheThrPheSerSer..... 31
112 ACTATGCTATGGCTGGTTCGCCAGGTTCCAGGGAAGAGCGTGAGTTT 161
|||||:||||| ||| |||||||||||||||:|||||
32 ..TyrAlaMetHisTrpValArgGlnAlaProGlyLysGlyLeuGluTrp 47
162 GTCCGACCTATTAGGTGGAGTGGTAAAGACATGGTATTAAGACTCGT 211
|||||:||||| :|||||:|||||: ||||| |||||||
48 ValAlaValIleSerTyrAspGlySerAsnLysTyrTyrAlaAspSerVa 64
212 GAAGGGCCGATTCCACCATCTCCAGAGATAAGCCCAAGACTACGCTTATC 261
|||||:||||| |||||||||||||||:|||||:|||||:|||||
64 LysGlyArgPheThrIleSerArgAspAsnSerLysAsnThrLeuTyrL 81
262 TGCAATGAACAGCTGAACACCTGAAGATACGGCGGTTTATTATTGTGCC 311
|||||:|||||:|||||:|||||:|||||:|||||:|||||
81 euGlnMetAsnSerLeuArgAlaGluAspThrAlaValTyrTyrCysAla 97
312 GCTCGACCGGTCCCGTGGATGATATTTCCCTGCCGTTGGGTTTGACTA 361
||| :|||||:
98 ArgArgSerAlaArgThr.....TyrTyrPheAspTy 108
362 CTGGGGCCAGGGACCCAGGTCACCGTCTCCTCA 395
|||||:|||||:|||||:|||||:|||||
108 rTrpGlyGlnGlyThrLeuValThrValSerSer 119
```

/SDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.AAY28615	+	421.00	716.94	7.1e-32
/SDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAB62215	+	421.00	716.63	7.1e-32
/SDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAE10561	+	421.00	716.56	7.1e-32
/SDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.AAY41173	+	420.00	710.39	9.9e-32
/SDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AAB72413	+	420.00	701.81	1.2e-31
/SDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.AAY79102				

```

Command line parameters:
-MODEL=frame_n2p_model -DEV=slp
-o*/cpn2.1/USPTO.spool/US09742690/runat_12082002_094914_76/a
-DB=A_Genesec_032802 -QFMT=fastan -SUFFIX=rsg -GAPOF=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000
-QGAPOF=4.500 -QGAPEXT=0.050 -XGAPOF=10.000 -XGAPEXT=0.500
-FCGAPOF=6.000 -FCGAPEXT=7.000 -YGAPOF=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blossum62
-TRANS=human40.cdi -LIST=45 -DICALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09742690.@Cn1_1.87 -NCPU=6 -ICPU=3 -LONGLOG
-DEV TIMEOUT=1200 -WARN TIMEOUT=30 -NO XLXPY -WAIT_THREADS=1

```

Search information block:

```

Search information block:
Query: US-09-742-690-1
Query length: 558
Database: A_Geneseq_032802:*
Database sequences: 747574
Database length: 111073796
Search time (sec): 85.160000

```

score_list:

Sequence	Strd Orig	ZScore	Score Len	Documentation
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA791902	+	682.00	1163.22	9.5e-57
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05281	+	682.00	1162.42	9.7e-57
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05286	+	679.00	1147.69	2.4e-56
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05287	+	679.00	1147.54	2.4e-56
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA791903	+	570.00	971.58	4.5e-46
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA791908	+	553.00	942.49	1.9e-44
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	550.50	931.77	3.8e-44
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	544.00	927.09	1.3e-43
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA791904	+	541.00	915.99	3.0e-43
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791904	+	541.00	915.95	3.0e-43
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791904	+	539.00	912.65	4.7e-43
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791904	+	517.00	879.56	5.2e-41
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	517.00	869.71	6.6e-41
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	508.50	865.08	3.4e-40
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	508.50	855.19	4.3e-40
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05284	+	473.00	804.08	8.2e-37
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05285	+	473.00	803.79	8.2e-37
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05288	+	471.00	800.72	1.3e-36
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	468.50	798.13	2.1e-36
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	449.00	765.32	1.5e-34
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	447.00	761.27	2.4e-34
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	446.50	760.80	2.6e-34
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05283	+	443.50	753.48	5.3e-34
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05283	+	443.00	754.27	7.5e-34
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05283	+	442.50	753.72	6.3e-34
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	441.00	744.80	1.0e-33
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05283	+	441.00	740.86	9.6e-34
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	439.50	749.15	1.2e-33
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	439.50	749.15	1.2e-33
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	438.50	746.80	1.5e-33
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	438.00	746.02	1.7e-33
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05282	+	433.00	735.45	5.3e-33
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05282	+	429.50	731.02	1.1e-32
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAE05282	+	429.00	717.21	1.7e-32
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2000.DAT:AAE05282	+	426.50	721.37	2.4e-32
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	425.50	724.18	2.7e-32
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05282	+	424.00	721.69	3.7e-32
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA791908	+	422.00	719.63	5.5e-32
SIDS5/cgcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05282	+	421.00	717.11	7.0e-32

seq documentation block:

seq_documentation_block:
ID AAY79102 standard: protein: 128 AA:

AA
AC
AA79102;

23-MAY-2000 (first entry)

XX llama heavy chain variable region recombinant fragment R2.

Antibody: heavy chain variable region; protein secretion; XX KW

KW glycosylation: llama: *Saccharomyces cerevisiae*.

XX
QS
Tama quanicoe alama.

XX
PN
W0200005389-A2

XX
03-FEB-2000XX
PE
08-III-1999. 99W0-EP04919XX
PB 20-JUL-1998 08EP-020243ZXX
DA (INTL.) INTL.EVER NV.

PA (UNIL) UNILEVER NV:
PA (UNIL) UNILEVER PLC:

XX Frenken I.G.J. Saet C. Verkleij A.J. Verrips C.T;

XX
DR WPI: 2000-182710/16.

High level recombinant production of heterologous protein with increased secretory efficiency in lower eukaryotic cells, useful for producing lysozyme, cutinase, or antibodies -

XX
ps Example 3d. Page 37. 55pp: English

XX The present sequence is that of llama anti-Azo dye RR6 heavy chain
CC variable region (VH) recombinant protein R2. Anti-RR6 VH DNA was
CC amplified by PCR (see AAZ58499-91), ligated into yeast multicopy
CC integration vectors and expressed in *Saccharomyces cerevisiae*.
CC Anti-RR6 VH fragments R2, R5 and R7 were isolated and sequenced
CC (see AAY79102-04). Secretion of the fragments was improved by the
CC introduction at specific sites of N-glycosylation sites. The
CC invention generally relates to the high level recombinant production
CC of proteins, such as antibody fragments, in lower eukaryotic host
CC (yeast or mould) cells. At least one N-glycosylation site is
CC introduced into the protein at a location between the N-terminus
CC and first hydrophobic region. The site is selected to increase the
CC level of secretion without affecting the protein's functionality.

Sequence 128 AA:

alignment scores:

alignment_scores:		
Quality:	682.00	Length: 128
Ratio:	5.328	Gaps: 0
Percent Similarity:	100.000	Percent Identity: 100.000

alignment block:

```

alignment_block: ..
US-09-742-690-1 x AAY79102
Align seg 1/1 to: AAY79102 from: 1 to: 128

```

12 CAGTGCAGCTGCAGGAGTCAGGGGGAGGATTGGTGCAGGCTGGGGCTC 61
|||||
1 GlnValGlnLeuGlnGlnSerGlyGlyGlyLeuValGlnAlaGlyGlySe 17

CC a DNA sequence encoding a heavy chain immunoglobulin, where the DNA is
CC linked to promoters and provided with an additional sequence encoding a
CC peptide capable of targeting heavy chain immunoglobulin to a cellular
CC compartment. The method is used for producing a heavy chain
CC immunoglobulin or an active fragment or derivative, or a protein that is
CC functionally equivalent for increasing the pathogen resistance in a plant
CC or to modulate metabolism in a plant. Under some circumstances it may be
CC desirable to retain the antibody product with the plant rather than
CC extracting and isolating the product. In particular, edible selected
CC antigens may be used in a method of passively immunising an animal,
CC preferably human, against the antigen, e.g., pathogenic organisms. The
CC present sequence is llama (camelid) anti-RR6 (RR6 is an azo-dye)
CC antibody (designated as HCV33) heavy chain variable domain (VH) attached
CC with an peptide linker, myc peptide.
XX
SQ Sequence 139 AA;

alignment_scores:
Quality: 682.00 Length: 128
Ratio: 5.328 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-742-690-1 x AAE05281
Align seg 1/1 to: AAE05281 from: 1 to: 139

12 CAGTGTGAGCTGAGGAGTGGTGGAGGAGTGGTGGAGGCTGGGGGCTC 61
1 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyGlyse 17
62 TCTGAGACTCTCTGTGAGCTCGGAGCGGCCAGCTGCTGATGTC 111
17 rLeuArgLeuSerCysAlaAlaSerGlyArgAlaThrSerGlyHisGlyH 34
112 ACTATGTATGGTGGTTCGCCAGGTTCCAGGAGGAGCGGTGAGTTT 161
34 iStyGlyMetGlyTrpPheArgGlnValProGlyLysGluArgGluPhe 50
162 GTCCAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211
51 ValAlaAlaIleArgTrpSerGlyLysGluThrTrpTrpLysAspSerVa 67
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTATC 261
67 LysGlyArgPheThrIleSerArgAspAsnAlaLysThrValTyrL 84
262 TGCAAAATGACAGCTGAACCTGAAGATACGGCGGTTTATATTGTC 311
84 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrTyrCysAla 100
312 GCTGAGCGGTCCGGTGGATGATATTTCCCTGGCGGTGGTGGTGGTACTA 361
101 AlaArgProValArgValAspAlaSerLeuProValGlyPheAspTy 117
362 CTGGGGCCAGGGAGCCAGGTCACCGCTCCTCTCA 395
117 rTrpGlyGlnGlyThrGlnValThrValSerSer 128

seq_name: /STD55/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05286
seq_documentation_block:
ID AAE05286 standard; Protein: 377 AA.
XX
AC AAE05286;
XX
DT 18-SEP-2001 (first entry)
XX
DE Llama HCV33 VH region with heavy chain constant regions.
XX
KW Camelid; llama; heavy chain immunoglobulin; pathogen resistance;
KW metabolism modulator; passive immunisation; heavy chain variable domain;

62 TCTGAGACTCTCTGTGAGCTCGGAGCGGCCAGGCTGGTGTGATGTC 111
17 rLeuArgLeuSerCysAlaAlaSerGlyArgAlaThrSerGlyHisGlyH 34
112 ACTATGTATGGTGGTGGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 161
34 iStyGlyMetGlyTrpPheArgGlnValProGlyLysGluArgGluPhe 50
162 GTCGAGCTATTAGTGGAGTGGTAAAGAGACATGGTATAAGACTCCGT 211
51 ValAlaAlaIleArgTrpSerGlyLysGluThrTrpTrpLysAspSerVa 67
212 GAAGGGCCGATTCCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
67 LysGlyArgPheThrIleSerArgAspAsnAlaLysThrValTyrL 84
262 TGCAAAATGACAGCTGAACCTCAAGATACGGCGGTATTATTTGTC 311
84 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrTyrCysAla 100
312 GCTGAGCGGTCCGGTGGATGATATTTCCCTGGCGGTGGTGGTGGTACTA 361
101 AlaArgProValArgValAspAlaSerLeuProValGlyPheAspTy 117
362 CTGGGGCCAGGGAGCCAGGTCACCGCTCCTCTCA 395
117 rTrpGlyGlnGlyThrGlnValThrValSerSer 128

seq_name: /STD55/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05281
seq_documentation_block:
ID AAE05281 standard; Protein: 139 AA.
XX
AC AAE05281;
XX
DT 18-SEP-2001 (first entry)
XX
DE Llama HCV33 antibody heavy chain variable domain with myc tag.
XX
KW Camelid; llama; heavy chain immunoglobulin; pathogen resistance;
KW metabolism modulator; passive immunisation; heavy chain variable domain;
KW VH; anti-RR6 antibody; HCV33; azo-dye.
XX
OS Chimeric - Lama guanicoe llama.
OS Chimeric - Unidentified.
XX
PN EP1118669-A2.
XX
PD 25-JUL-2001.
XX
PF 08-DEC-2000; 2000EP-0310997.
XX
PR 17-DEC-1999; 99EP-0310188.
XX
PA (UNIL) UNILEVER PLC.
PA (UNIL) UNILEVER NV.
XX
PI Frenken LGJ, Van Der Logt CPE, Jobling SA, Teh Y;
XX
DR WPI: 2001-427157/46.
DR N-PSDB; AADI0052.

Modifying a plant to produce an antibody useful for increasing pathogen resistance or to modulate metabolism comprises introducing a DNA sequence encoding a heavy chain immunoglobulin linked to a peptide that targets a cellular compartment -
Example 1; Fig 1; 81pp; English.
The present invention relates to a method for modifying a plant to produce an antibody or an active fragment or derivative, or a protein functional equivalent, in a cellular compartment comprises introducing

KW VH; anti-RR6 antibody; HCV33; azo-dye.
 XX Chimeric - Lama guanicoe glama.
 OS Chimeric - Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Region 1..129
 FT /note= "Corresponds to Llama HVC33 VH region"
 FT Region 130..158
 FT /note= "Corresponds to hinge region"
 FT Region 159..268
 FT /note= "Corresponds to CH2 heavy chain constant region"
 FT Region 269..377
 FT /note= "Corresponds to CH3 heavy chain constant region"
 XX
 XX EP1118669-A2.
 PN
 PD 25-JUL-2001.
 XX
 PF 08-DEC-2000; 2000EP-0310997.
 XX
 PR 17-DEC-1999; 99EP-0310188.
 XX
 XX (UNIL) UNILEVER PLC.
 PA (UNIL) UNILEVER NV.
 XX
 XX Frenken LGJ, Van Der Logt CPE, Jobling SA, Teh Y;
 PI
 XX WPI: 2001-427157/46.
 DR N-PSDB; AADI0057.
 DR
 XX
 PT Modifying a plant to produce an antibody useful for increasing pathogen
 PT resistance or to modulate metabolism comprises introducing a DNA
 PT sequence encoding a heavy chain immunoglobulin linked to a peptide that
 PT targets a cellular compartment -
 XX
 PS Example 12; Fig 18; 81pp; English.
 CC
 CC The present invention relates to a method for modifying a plant to
 CC produce an antibody or an active fragment or derivative, or a protein
 CC functional equivalent, in a cellular compartment comprising introducing
 CC a DNA sequence encoding a heavy chain immunoglobulin, where the DNA is
 CC linked to promoters and provided with an additional sequence encoding a
 CC peptide capable of targeting heavy chain immunoglobulin to a cellular
 CC compartment. The method is used for producing a heavy chain
 CC immunoglobulin or an active fragment or derivative, or a protein that is
 CC functionally equivalent for increasing the pathogen resistance in a plant
 CC or to modulate metabolism in a plant. Under some circumstances it may be
 CC desirable to retain the antibody product with the plant rather than
 CC extracting and isolating the product. In particular, edible selected
 CC antigens may be used in a method of passively immunising an animal,
 CC preferably human, against the antigen, e.g., pathogenic organisms. The
 CC present sequence is a llama (camelid) anti-RR6 (RR6 is an azo-dye)
 CC antibody (designated as HCV33) heavy chain variable domain (VH)
 CC linked to the non-classical heavy chain constant regions denoted as
 CC hinge-CH2-CH3.
 XX
 SQ Sequence 377 AA;

alignment_scores:
 Quality: 679.00 Length: 128
 Ratio: 5.305 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 99.219
 alignment_block:
 US-09-742-690-1 x AAE05286 ..
 Align seg 1/1 to: AAE05286 from: 1 to: 377
 12 CAGTCGAGCTGCAGGAGTACAGGGGAGGATTGGTCAGGCTGGGGCTC 61
 :::
 2 GluValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyGlyse 18

62 TCTGAGACTCTCTGTGCAGCCTCGGACGCGCCAGCCAGTGGTGCATGTC 111
 |||||
 18 rLeuArgLeuSerCysAlaAlaSerGlyArgAlaThrSerGlyHisGlyH 35
 |||||
 112 ACTATGGTATGGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTACGTTT 161
 |||||
 35 iStyGlyMetGlyTrpPheArgGlnValProGlyLysGluArgGluPhe 51
 |||||
 162 GTCCGACGCTATTAGTGGAGTGGTAAAGACATGGTATAAAGACTCCGT 211
 |||||
 52 ValAlaAlaIleArgTrpSerGlyLysGluThrTrpTyrLysAspSerVa 68
 |||||
 212 GAAGGGCCGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
 |||||
 68 LysGlyArgPheThrIleSerArgAspAsnAlaLysThrValTyrL 85
 |||||
 262 TGCAGATGACAGCCTGAACCTGAAGATACGGCGCTTATTATTGTGCC 311
 |||||
 85 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrTyrCysAla 101
 |||||
 312 GCTGACCGGTCGCGTGATGATATTCCCTGCCGTTGGTTTGACTA 361
 |||||
 102 AlaArgProValArgValAspIleSerLeuProValGlyPheAspTy 118
 |||||
 362 CTGGGGCCAGGGACCCAGGTCCCGTCTCTCTCA 395
 |||||
 118 rTrpGlyGlnGlyThrGlnValThrValSerSer 129
 |||||
 seq_name: /SDS5/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:AAE05287
 seq_documentation_block:
 ID AAE05287 standard; Protein; 383 AA.
 XX
 AC AAE05287;
 XX
 DT 18-SEP-2001 (first entry)
 XX
 DE Llama HCV33 VH region with ER retention signal and hinge-CH2-CH3.
 XX
 KW Camelid; llama; heavy chain immunoglobulin; pathogen resistance;
 KW metabolism modulator; passive immunisation; heavy chain variable domain;
 KW VH; anti-RR6 antibody; HCV33; azo-dye.
 XX
 OS Chimeric - Lama guanicoe glama.
 OS Chimeric - Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Region 1..129
 FT /note= "Corresponds to Llama HVC33 VH region"
 FT Region 130..158
 FT /note= "Corresponds to hinge region"
 FT Region 159..268
 FT /note= "Corresponds to CH2 heavy chain constant region"
 FT Region 269..377
 FT /note= "Corresponds to CH3 heavy chain constant region"
 FT Region 378..383
 FT /note= "Corresponds to ER retention signal sequence"
 XX
 XX EP1118669-A2.
 XX
 PD 25-JUL-2001.
 XX
 PF 08-DEC-2000; 2000EP-0310997.
 XX
 PR 17-DEC-1999; 99EP-0310188.
 XX
 XX (UNIL) UNILEVER PLC.
 PA (UNIL) UNILEVER NV.
 XX
 XX Frenken LGJ, Van Der Logt CPE, Jobling SA, Teh Y;
 XX
 XX WPI: 2001-427157/46.

DR N-PSDB; AAD10058.
XX
PT Modifying a plant to produce an antibody useful for increasing pathogen
PT resistance or to modulate metabolism comprises introducing a DNA
PT sequence encoding a heavy chain immunoglobulin linked to a peptide that
PT targets a cellular compartment -
XX
XX
PS Example 12; Fig 19; 81pp; English.
XX
CC The present invention relates to a method for modifying a plant to
CC produce an antibody or an active fragment or derivative, or a protein
CC functional equivalent, in a cellular compartment comprising introducing
CC a DNA sequence encoding a heavy chain immunoglobulin, where the DNA is
CC linked to promoters and provided with an additional sequence encoding a
CC peptide capable of targeting heavy chain immunoglobulin to a cellular
CC compartment. The method is used for producing a heavy chain
CC immunoglobulin or an active fragment or derivative, or a protein that is
CC functionally equivalent for increasing the pathogen resistance in a plant
CC or to modulate metabolism in a plant. Under some circumstances it may be
CC desirable to retain the antibody product with the plant rather than
CC extracting and isolating the product. In particular, edible selected
CC antigens may be used in a method of passively immunising an animal,
CC preferably human, against the antigen, e.g., pathogenic organisms. The
CC present sequence is a llama (camelid) anti-RR6 (RR6 is an azo-dye)
CC antibody (designated as HC933) heavy chain variable domain (VH) linked
CC to the non-classical heavy chain constant regions denoted as
CC hinge-CH2-CH3 and endoplasmic reticulum (ER) retention signal.
XX
XX
SQ Sequence 383 AA;

alignment_scores:
Quality: 679.00 Length: 128
Ratio: 5.305 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 99.219
alignment_block:
US-09-742-690-1 x AAE05287 ..
Align seg 1/1 to: AAE05287 from: 1 to: 383
12 CAGTGCAGCTGAGGAGTGCAGGGGAGGATTGGTCAGGCTGGGGCTC 61
:::|||||
2 GluValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyGly 18
18 rLeuArgLeuSerCysAlaAlaSerGlyArgAlaThrSerGlyHisGlyH 35
62 TCTGAGACTCTCTGTGACGCTCGGACGCGCCACGAGTGGTCATGCTC 111
|||||
112 ACTATGATGGCTGGTTCGCCAGGTTCCAGGAGAGCGTGAGTTT 161
|||||
35 IStyrGlyMetGlyTrpPheArgGlnValProGlyLysGluArgGluPhe 51
162 GTCGACGATATTAGGTGGTGTAAAGACATGGTATAAAGACTCCGT 211
|||||
52 ValAlaAlaIleArgTrpSerGlyLysGluThrTrpTrpLysAspSera 68
212 GAAGGGCCGATTCCACCTCCAGAGATAACGCCAGACTACGTTTATC 261
|||||
68 LysGlyArgPheThrIleSerArgAspAsnAlaLysThrValTyrIL 85
262 TGCAAATGAACACCTGAACCTCAAGATACGGCCGTTTATTATTGTGC 311
|||||
85 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrCysAla 101
312 GCTGGACCGTCCGGTGGATGATATTTCCTCCCGTGGGTGGTGA 361
102 AlaArgProValArgValAspAspIleSerLeuProValGlyPheAsp 118
362 CTGGGGCCAGGGACCCAGGTACCGCTCCTCTCA 395
118 rTrpGlyGlnGlyThrGlnValThrValSerSer 129

seq_name: /STDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA79103
seq_documentation_block:
ID AAY79103 standard; Protein; 128 AA.
XX
AC AAY79103;
XX
DT 23-MAY-2000 (first entry)
XX
DE Llama heavy chain variable region recombinant fragment R5.
XX
KW Antibody; heavy chain variable region; protein secretion;
KW glycosylation; llama; Saccharomyces cerevisiae.
XX
OS Lama guanicoe glama.
XX
PN WO200005389-A2.
XX
PD 03-FEB-2000.
XX
PF 08-JUL-1999; 99WO-EF04919.
XX
PR 20-JUL-1998; 98EP-0202432.
XX
PA (UNIL) UNILEVER NV.
PA (UNIL) UNILEVER PLC.
XX
PI Frenken LGJ, Sajt C, Verkleij AJ, Verrips CT;
XX
DR WPI; 2000-182710/16.
XX
PT High level recombinant production of heterologous protein with
PT increased secretory efficiency in lower eukaryotic cells, useful for
PT producing lysozyme, cutinase, or antibodies -
PS Example 3d; Page 37; 55pp; English.
XX
CC The present sequence is that of llama anti-Azo dye RR6 heavy chain
CC variable region (VH) recombinant protein R5. Anti-RR6 VH DNA was
CC amplified by PCR (see AA258489-91), ligated into yeast multicopy
CC integration vectors and expressed in Saccharomyces cerevisiae.
CC Anti-RR6 VH fragments R2, R5 and R7 were isolated and sequenced
CC (see AA79102-04). Secretion of the fragments was improved by the
CC introduction at specific sites of N-glycosylation sites. The
CC invention generally relates to the high level recombinant production
CC of proteins, such as antibody fragments, in lower eukaryotic host
CC (yeast or mould) cells. At least one N-glycosylation site is
CC introduced into the protein at a location between the N-terminus
CC and first hydrophobic region. The site is selected to increase the
CC level of secretion without affecting the protein's functionality.
XX
SQ Sequence 128 AA;

alignment_scores:
Quality: 570.00 Length: 128
Ratio: 4.790 Gaps: 0
Percent Similarity: 92.969 Percent Identity: 85.938
alignment_block:
US-09-742-690-1 x AAY79103 ..
Align seg 1/1 to: AAY79103 from: 1 to: 128
12 CAGTGCAGCTGAGGAGTGCAGGGGAGGATTGGTCAGGCTGGGGCTC 61
|||||
1 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyGly 17
62 TCTGAGACTCTCTGTGACGCTCGGACGCGCCACGAGTGGTCATGCTC 111
|||||
17 rLeuArgLeuSerCysAlaAlaSerGlyArgThrSerHisGlyTyrGlyG 34
112 ACTATGATGGCTGGTTCGCCAGGATTCCAGGAGAGCGTGAGTTT 161

PI Frenken LGJ, Howell S, Ledeboer AM, Van Der Logt CPE;
 XX WPI: 1999-313342/26.
 DR N-PSDB; AAX59616.
 XX
 XX Multivalent antigen binding protein useful for inactivation of
 XX (bacteriophages or viruses
 XX
 XX Disclosure; Fig 12; 89pp; English.
 XX
 CC The specification describes multivalent antigen binding proteins
 CC comprising a single polypeptide chain comprising, in series, two or
 CC more single domain binding sites. The multivalent antigen proteins
 CC are useful in applications where antibodies have been used in the
 CC prior art, including diagnosis, therapy, targeting, immunoassays,
 CC cross-linking methods including agglutination, or for purification
 CC processes. It is also useful for inactivation of (bacteriophages
 CC or viruses, and in detergents. The present sequence represents an
 CC homodimeric bivalent anti-RR6 antigen binding protein.
 XX
 XX Sequence 250 AA;

alignment_scores:
 Quality: 550.50 Length: 161
 Ratio: 4.301 Gaps: 1
 Percent Similarity: 79.503 Percent Identity: 68.944

alignment_block:

US-09-742-690-1 x AAY15413 ..

Align seg 1/1 to: AAY15413 from: 1 to: 250

21 CTGCGAGGAGTACGGGGAGGAGTGTGTGAGGCTGGGGCTCTCTGAGACT 70
 1 LeuGlnLuserGlyGlyLeuValGlnAlaGlyAspSerLeuArgLe 17
 71 CTCCTGTGAGCCTCGGGACGGCCACGAGTGGTTCATGCTCATGTGTA 120
 17 uSerCysAlaAlaSerGlyArgThrSerHisGlyTyrGlyTyrGlyM 34
 121 TGGGCTGGTTCCTCCAGGTCCTCCAGGAGGAGGCGTTCAGTTTCTCCAGCT 170
 34 eGlyTyrPheArgGlnLeuProGlyLysGluArgGluLeuAlaAla 50
 171 ATTAGTGAGTGGTAAAGACATGGTATTAAGACTCCGTGAAGGGCG 220
 51 IleArgTyrSerGlyArgAsnThrTyrTyrAlaAspSerValLysGlyAr 67
 221 ATTACCAATCTCCAGAGATAACGCCAAGACTACGGTTTATCTGCAATGA 270
 67 gPheThrIleSerArgAspAsnValLysAspMetLeuTyrLeuGlnMeta 84
 271 ACAGCTGAACCTCAAGATACGGCGTTTATTATTGTCCGCTGACCG 320
 84 snSerLeuLysProGluAspThrAlaValTyrThrCysAlaValArgThr 100
 321 GTCGCGTGGATGATATTCCTCCGCGTGGTGGTGGTGGTGGTGGTGGTGGT 370
 101 ValArgValValAspIleSerSerProValGlyPheAlaTyrTrpGlyGl 117
 371 GGGGACCGAGTCCCGCTCTCTAGGATCTCATCACCATCACCATCAGG 420
 117 nglyThrGlnValThrValSerSerGlnValGlnLeuGlnLuserGlyG 134
 421 GA.....TCCACCTCCATTGAA 437
 134 lyGlyLeuValGlnAlaGlyAspSerLeuArgLeuSerCysAlaAlaSer 150
 438 GGTCTGACCGAGTCTCACTACAGCTCAGTGTGGT 470
 151 GlyArgThrSerHisGlyTyrGlyGlyTyrGly 161

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:AA79104
 seq_documentation_block:
 ID AAY79104 standard; Protein: 128 AA.

XX
 AC AAY79104;
 XX
 XX 23-MAY-2000 (first entry)
 DT
 XX
 DE Llama heavy chain variable region recombinant fragment R7.
 XX
 KW Antibody; heavy chain variable region; protein secretion;
 KW glycosylation; llama; Saccharomyces cerevisiae.
 XX
 OS Lama guanicoe glama.
 XX
 PN W0200005389-A2.
 XX
 XX 03-FEB-2000.
 PD
 XX
 XX 08-JUL-1999; 99WO-EP04919.
 PF
 XX
 XX 20-JUL-1998; 98EP-0202432.
 PR
 XX (UNIL) UNILEVER NV.
 PA (UNIL) UNILEVER PLC.
 XX
 XX Frenken LGJ, Sagt C, Verkleij AJ, Verrips CT;
 XX
 DR WPI: 2000-182710/16.
 XX

High level recombinant production of heterologous protein with increased secretory efficiency in lower eukaryotic cells, useful for producing lysozyme, cutinase, or antibodies -
 Example 3d; Page 37; 55pp; English.

The present sequence is that of llama anti-Azo dye RR6 heavy chain variable region (VH) recombinant protein R7. Anti-RR6 VH DNA was amplified by PCR (see AAZ58489-91), ligated into Yeast multicopy integration vectors and expressed in Saccharomyces cerevisiae. Anti-RR6 VH fragments R2, R5 and R7 were isolated and sequenced (see AAY79102-04). Secretion of the fragments was improved by the introduction at specific sites of N-glycosylation sites. The invention generally relates to the high level recombinant production of proteins, such as antibody fragments, in lower eukaryotic host (yeast or mould) cells. At least one N-glycosylation site is introduced into the protein at a location between the N-terminus and first hydrophobic region. The site is selected to increase the level of secretion without affecting the protein's functionality.

SQ Sequence 128 AA;

alignment_scores:
 Quality: 544.00 Length: 128
 Ratio: 4.650 Gaps: 0
 Percent Similarity: 91.406 Percent Identity: 81.250

alignment_block:

US-09-742-690-1 x AAY79104 ..

Align seg 1/1 to: AAY79104 from: 1 to: 128

12 CAGGTGCAGTGCAGGAGTGCAGGGGAGGATTGGTGCAGGCTGGGGGCTC 61
 1 GlnValGlnLeuGlnGluSerGlyGlyLeuValGlnThrGlyAspSe 17
 62 TCTGAGACTCTCTCTGTCAGCCTCGGAGCGCCGCCACCTGGTGTATGGTC 111
 17 rLeuArgLeuSerCysAlaAlaSerGlyArgThrSerHisGlyTyrGlyG 34
 112 ACTATGGTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGAGTTT 161


```

|||||
34  lyTyrglyMetGlyTrpPheArgGlnIleProGlyLysGluArgGluLeu 50
162 GTCCGACGATTAGTGGAGTGTAAAGACATGCTATATAAGACTCCGT 211
|||||
51  ValAlaAlaIleArgTrpSerGlyArgAsnThrTyrrAlaAspSerVa 67
212 GAAGGCCGATTACCATCTCCAGAGATACGCCAAGACTACGGTTATC 261
|||||
67  llysglyArgPheThrIleSerArgAspAsnValLysAspMetLeuTyrl 84
262 TGCAATGAACAGCTGAACCTGAAGATACGCCGCTTTATTATTGTGCC 311
|||||
84  euGlnMetAspSerLeuLysProGluAspThrAlaValThrTyrrCysAla 100
312 GCTCGACCGCTCCGCGTGGATGATATTCCTCCCGGTTGGGTTTGACTA 361
|||||
101 ValArgThrValArgValValAspIleSerSerProValGlyPheAlaty 117
362 CTGGGCCCGGGGACCCAGGTCACTCCCTCA 395
117 rTrpGlyGlnGlyThrGlnValThrValSerSer 128

```

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA15406

seq_documentation_block:

ID AAY15406 standard; Protein; 238 AA.

XX AAY15406;

DT 21-JUL-1999 (first entry)

XX Aanti-hCG anti-RR6 antigen binding protein.

KW Multivalent antigen binding protein; single polypeptide chain;
 KW single domain binding site; diagnosis; therapy; targeting;
 KW immunoassay; cross-linking; agglutination; purification;
 KW phage inactivation; detergent.

XX Synthetic.

XX WO9923221-A2.

XX 14-MAY-1999.

XX 27-OCT-1998; 98WO-EP06991.

XX 27-OCT-1997; 97EP-0308538.

XX (UNIL) UNILEVER NV.

XX (UNIL) UNILEVER PLC.

XX Frenken IGJ, Howell S, Ledebor AM, Van Der Logt CPE;

XX WPI; 1999-313342/26.

XX N-PSDB; AAX59612.

XX Multivalent antigen binding protein useful for inactivation of
 (bacteriophages or viruses

PS Disclosure; Fig 27; 89pp; English.

XX The specification describes multivalent antigen binding proteins
 CC comprising a single polypeptide chain comprising, in series, two or
 CC more single domain binding sites. The multivalent antigen proteins
 CC are useful in applications where antibodies have been used in the
 CC prior art, including diagnosis, therapy, targeting, immunoassays,
 CC cross-linking methods including agglutination, or for purification
 CC processes. It is also useful for inactivation of (bacteriophages
 CC or viruses, and in detergents. The present sequence represents an
 CC anti-hCG anti-RR6 bispecific biheaded antigen binding protein.

XX Sequence 238 AA;

alignment_scores:

Quality: 541.00 Length: 125
 Ratio: 4.746 Gaps: 0
 Percent Similarity: 91.200 Percent Identity: 82.400

alignment_block:

US-09-742-690-1 x AAY15406 ..

Align seg 1/1 to: AAY15406 from: 1 to: 238

```

12 CAGGTCACGCTCCAGGAGTCCAGGGGAGGATTGTCAGGCTGGGGCTC 61
|||||
114 GlnValGlnLeuGlnSerGlyGlyLeuValGlnAlaGlyAspSe 130
62 TCTGAGACTCTCTGTGCGACCTCGGGACCGCCAGCCAGTGGTCATGTC 111
|||||
130 rLeuArgLeuSerCysAlaAlaSerGlyArgThrSerHisGlyTyrrGlyG 147
112 ACTATGTTATGGCTGGTCCCGCAGGTTCCAGGGAAGAGCGGTGACTTT 161
|||||
147 lyTyrglyMetGlyTrpPheArgGlnIleProGlyLysGluArgGluLeu 163
162 GTCGCGACTATTAGTGGAGTGTAAAGACATGTTATTAAGACTCCGT 211
|||||
164 ValAlaAlaIleArgTrpSerGlyArgAsnThrTyrrAlaAspSerVa 180
212 GAAGGCCGATTACCATCTCCAGATACGCCAAGACTACGGTTTATC 261
|||||
180 llysglyArgPheThrIleSerArgAspAsnValLysAspMetLeuTyrl 197
262 TGCAATGAACAGCTGAACCTGAAGATACGCCGCTTTATTATTGTGCC 311
|||||
197 euGlnMetAsnSerLeuLysProGluAspThrAlaValThrTyrrCysAla 213
312 GCTCGACCGCTCCGCGTGGATGATATTCCTCCCGGTTGGGTTTGACTA 361
|||||
214 ValArgThrValArgValValAspIleSerSerProValGlyPheAlaty 230
362 CTGGGCCCGGGGACCCAGGTCACTCC 386
230 rTrpGlyGlnGlyThrGlnValThr 238

```

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA15408

seq_documentation_block:

ID AAY15408 standard; Protein; 239 AA.

XX AAY15408;

DT 21-JUL-1999 (first entry)

XX Anti-hCG anti-RR6 antigen binding protein.

KW Multivalent antigen binding protein; single polypeptide chain;
 KW single domain binding site; diagnosis; therapy; targeting;
 KW immunoassay; cross-linking; agglutination; purification;
 KW phage inactivation; detergent.

XX Synthetic.

XX WO9923221-A2.

XX 14-MAY-1999.

XX 27-OCT-1998; 98WO-EP06991.

XX 27-OCT-1997; 97EP-0308538.

XX (UNIL) UNILEVER NV.

XX (UNIL) UNILEVER PLC.

XX

PI Franken LGJ, Howell S, Ledeboer AM, Van Der Logt CPE;
XX WPI; 1999-313342/26.
DR N-PSDB; AAX59614.
XX Multivalent antigen binding protein useful for inactivation of
PT (bacteriophages or viruses
XX Disclosure; Fig 10; 89pp; English.
XX The specification describes multivalent antigen binding proteins
CC comprising a single polypeptide chain comprising, in series, two or
CC more single domain binding sites. The multivalent antigen proteins
CC are useful in applications where antibodies have been used in the
CC prior art, including diagnosis, therapy, targeting, immunoassays,
CC cross-linking methods including agglutination, or for purification
CC processes. It is also useful for inactivation of (bacteriophages
CC or viruses, and in detergents. The present sequence represents an
CC anti-HCG anti-RR6 bispecific biheaded antigen binding protein (ABP).
XX
XX Sequence 239 AA;

alignment_scores:
Quality: 541.00 Length: 125
Ratio: 4.746 Gaps: 0
Percent Similarity: 91.200 Percent Identity: 82.400

alignment_block:
US-09-742-690-1 x AAY15408 ..

Align seg 1/1 to: AAY15408 from: 1 to: 239

12 CAGGTGCAGCTCAGGAGTGCAGGGAGGATGTGTCAGGCTGGGGCTC 61
115 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyAspSe 131
62 TCTGAGACTCTCCTGTGCAGCTCGGGAGCGGCCACAGTGTCTATGTC 111
131 rLeuArgLeuSerCysAlaAlaSerGlyArgThrHisGlyTyrGlyG 148
112 ACTATGTTGATGGCTGTCCTCCAGGTCCAGGAGGAGGCGTGTGTT 161
148 lyTyrGlyMetGlyTrpPheArgGlnIleProGlyLysGluArgGluLeu 164
162 GTCCGAGCTATTAGTGTGAGTGGTAAAGAGACATGGTATAGACTCCGT 211
165 ValAlaAlaIleArgTrpSerGlyArgAsnThrTyrTyrAlaAspSerVa 181
212 GAAGGCGCGATTTCACATCTCCAGAGATAACGCCAAGACTACGGTTTATC 261
181 lLysGlyArgPheThrIleSerArgAspAsnValLysAspMetLeuTyrL 198
262 TCCAATGACAGCCTGAACCTGAGATACGGCGGTTTATATTGTGCC 311
198 euGlnMetAsnSerLeuLysProGluAspThrAlaValTyrThrCysAla 214
312 GCTCAGCGGTCGGCTGATGATATTCCCTGCGGTTGGGTTTGACTA 361
215 ValArgThrValArgValValAspIleSerSerProValGlyPheAlaTy 231
362 CTGGGCGCAGGGGCCAGGTCACC 386
231 rTrpGlyGlnGlyThrGlnValThr 239

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: AAY15414
seq_documentation_block:
ID AAY15414 standard; Protein: 236 AA.
XX
AC AAY15414;
XX
DT 21-JUL-1999 (first entry)

XX An anti-RR6 antigen binding protein.
DE
XX Multivalent antigen binding protein; single polypeptide chain;
KW single domain binding site; diagnosis; therapy; targeting;
KW immunoassay; cross-linking; agglutination; purification;
XX phage inactivation; detergent.
XX
OS Synthetic.
XX
PN WO9923221-A2.
XX
XX 14-MAY-1999.
XX
XX 27-OCT-1998; 98WO-EP06991.
XX
XX 27-OCT-1997; 97EP-0308538.
XX
XX (UNIL) UNILEVER NV.
PA (UNIL) UNILEVER PLC.
XX
XX Franken LGJ, Howell S, Ledeboer AM, Van Der Logt CPE;
XX
XX WPI; 1999-313342/26.
DR N-PSDB; AAX59617.
XX
XX Multivalent antigen binding protein useful for inactivation of
PT (bacteriophages or viruses
XX Disclosure; Fig 13; 89pp; English.
XX
XX The specification describes multivalent antigen binding proteins
CC comprising a single polypeptide chain comprising, in series, two or
CC more single domain binding sites. The multivalent antigen proteins
CC are useful in applications where antibodies have been used in the
CC prior art, including diagnosis, therapy, targeting, immunoassays,
CC cross-linking methods including agglutination, or for purification
CC processes. It is also useful for inactivation of (bacteriophages
CC or viruses, and in detergents. The present sequence represents an
CC heterodimeric bivalent anti-RR6 antigen binding protein.
XX
XX Sequence 236 AA;

alignment_scores:
Quality: 539.00 Length: 125
Ratio: 4.728 Gaps: 0
Percent Similarity: 91.200 Percent Identity: 82.400

alignment_block:
US-09-742-690-1 x AAY15414 ..

Align seg 1/1 to: AAY15414 from: 1 to: 236

21 CTGAGGAGTACAGGGGAGGATGTGTCAGGCTGGGGCTCTCTGAGACT 70
1 LeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyAspSerLeuArgLe 17
71 CTCCTGTGCAGCTCGGAGCGGCCACAGTGTGTCATGTCATCTATGTA 120
17 uSerCysAlaAlaSerGlyArgThrSerHisGlyTyrGlyTyrGlyM 34
121 TGGGCTGTGTCGCCAGGTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGT 170
34 etGlyTrpPheArgGlnIleProGlyLysGluArgGluLeuValAlaAla 50
171 ATTAGGTGGTGGTAAAGAGACATGGTATAAAGACTCCGTAAGGCGCG 220
51 lleArgTrpSerGlyArgAsnThrTyrTyrAlaAspSerValLysGlyAr 67
221 ATTACACTCTCCAGAGATACGCCAAGACTACGGTTTATCTGCAATGA 270
67 gPheThrIleSerArgAspAsnValLysAspMetLeuTyrLeuGlnMeta 84

271 ACAGCTGAACCTGAGATACGGCCGTTATTATTGTGCGCTGACCG 320
|||||
84 snSerLeuLysProGluAspThrAlaValTyrThrCysAlaValArgThr 100
|||||
321 CTCGCGGGGATGATATTCCTCCGCGTTGGTTGACTACTGGGCCA 370
|||||
101 ValArgValAlaAspIleSerProValGlyPheAlaTyrTrpGlyGI 117
|||||
371 GGGGACCCAGGTCACCGTCTCTCA 395
|||||
117 nGlyThrGlnValThrValSerSer 125
|||||

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA08986

seq_documentation_block:
ID_AAY08986 standard; Protein; 147 AA.
AC_AAY08986;
XX DT 20-AUG-1999 (first entry)
XX DE Llama protein construct pPIC.HCV3-His2t.
XX KW Analytical device; electrically conductive solid support; immobilisation;
KW electroactive; analysis; peptide hormone; steroid; disease marker; HCV3;
KW diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama.
OS Lama guanicoe glama.
OS Synthetic.
XX WO9927356-A1.
XX 03-JUN-1999.
XX 23-NOV-1998; 98WO-GB03495.
XX 21-NOV-1997; 97EP-0309425.
XX (UNIL) UNILEVER NV.
XX (UNIL) UNILEVER PLC.
XX Badley RA, Porter RA;
XX WPI; 1999-385228/32.
XX Analytical device including electrochemically active compound on
XX solid support

Example 6.2; Page 67; 78pp; English.

This invention describes a novel component for an analytical device comprising an electrically conductive solid support on which a compound (I) is immobilized which has an electroactive part (EA) with an electrochemical property that can be modulated in a detectable manner by binding it to a specific binding partner (II). Compounds used in the method of the invention include those with the formula R = hydrogen, hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally substituted), halo, amido, or amino; optionally one or more positions on the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or alkoxy (all optionally substituted), acid groups (organic or inorganic), halo, amido or amino. Devices that contain the component of the invention are used for qualitative and quantitative analysis of e.g. nucleic acid, hormones (peptide or steroid), disease markers, diagnostic indicators etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices can be used with turbid solutions (whole blood, milk, culture liquids etc.), unlike conventional colourimetric methods. This sequence represents a protein construct pPIC.HCV3-His2t which contains the llama HC-V fragment, HCV3 which is used in the method of the invention.

Sequence 147 AA;

alignment_scores:
Quality: 517.00 Length: 141
Ratio: 4.419 Gaps: 2
Percent Similarity: 82.979 Percent identity: 73.050
alignment_block:
US-09-742-690-1 x AAY08986
Align seg 1/1 to: AAY08986 from: 1 to: 147
12 CAGGTGCAGTCGAGGAGTTCAGGGGAGGATTGGTGCAGGCTGGGGCTC 61
|||||
1 GlnValGlnLeuGlnGluSerGlyGlyLeuValGlnProGlyGlySe 17
62 TCTGAGACTCTCTGTGTCAGCTCGGACGCCGCCACGAGTGTGTCATGCTC 111
|||||
17 rLeuArgLeuSerCysAlaAlaSerGlyLeuThrLeuThr.....T 31
112 ACTATGTTATGGCTGGTTCGCCAGGTTCCAGGGAAGGAGCGTGCAGTTT 161
|||||
31 hrTyrSerThrGlyTrpPheArgGlnAlaProGlyLysGluArgGluPhe 47
162 CTCGCAGCTATTAGTGGAGT...GGTAAAGAGACATGCTATAAAGACTC 208
|||||
48 ValGlyMetLeuGlyTrpSerGlyGlyGlyAsnThrTyrTyrAlaAspSe 64
209 CGTGAAGGCGGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTTT 258
|||||
64 rValLysGlyArgPheThrIleSerArgAspAsnAlaLysAsnMetValp 81
259 ATCTGCAATGAACAGCCTGAACCTGAACCTGAACCTGAACCTGTTATTATCT 308
|||||
81 heLeuGlnMetSerSerLeuLysProGluAspThrAlaValTyrCys 97
309 GCGCTCGACCGTCCGCGTGATGATATTCCTCCGCGGTTGGGTTTGA 358
|||||
98 AlaAlaArgGlnProTyrArgGlySerTyrSerAspProAsnAsnTyrHi 114
359 CTACTGGGGCCAGGGACCCAGGTCACCGTCTCCTCAGGATCTCATCACC 408
|||||
114 styTrpGlyGlnGlyThrGlnValThrValSerSerGlySerHisHisH 131
409 ATCACCATCAGGATCCACCTCC 431
|||||
131 IshiShiShiGlySerGlySer 138

seq_name: /SIDS5/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA08989

seq_documentation_block:
ID_AAY08989 standard; Protein; 409 AA.

XX AC_AAY08989;
XX DT 20-AUG-1999 (first entry)
XX DE Llama protein construct pPIC.scFv4155-link-HCV3-His2t.
XX KW Analytical device; electrically conductive solid support; immobilisation;
KW electroactive; analysis; peptide hormone; steroid; disease marker; HCV3;
KW diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama.
OS Lama guanicoe glama.
OS Synthetic.
XX WO9927356-A1.
XX 03-JUN-1999.
XX 23-NOV-1998; 98WO-GB03495.
XX 21-NOV-1997; 97EP-0309425.
XX (UNIL) UNILEVER NV.

PA (UNIL) UNILEVER PLC.
 XX Badley RA, Porter RA;
 XX WPI; 1999-385228/32.
 XX Analytical device including electrochemically active compound on
 PT solid support
 PT
 XX Example 6.4; Page 70-71; 78pp; English.
 XX This invention describes a novel component for an analytical device
 CC comprising an electrically conductive solid support on which a compound
 CC (I) is immobilized which has an electroactive part (EA) with an
 CC electrochemical property that can be modulated in a detectable manner by
 CC binding it to a specific binding partner (II). Compounds used in the
 CC method of the invention include those with the formula R = hydrogen,
 CC hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally
 CC substituted), halo, amido, or amino; optionally one or more positions on
 CC the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or
 CC alkoxy (all optionally substituted), acid groups (organic or inorganic),
 CC halo, amido or amino. Devices that contain the component of the invention
 CC are used for qualitative and quantitative analysis of e.g. nucleic acid,
 CC hormones (peptide or steroid), disease markers, diagnostic indicators
 CC etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices
 CC can be used with turbid solutions (whole blood, milk, culture liquids
 CC etc.), unlike conventional colourimetric methods. This sequence
 CC represents a protein construct pPIC.scFv4155-link-HCV3-His2t which
 CC contains the llama HC-v fragment, HCV3 which is used in the method of
 CC the invention.
 XX
 XX SQ Sequence 409 AA;

alignment_scores:
 Quality: 517.00 Length: 141
 Ratio: 4.419 Gaps: 2
 Percent Similarity: 82.979 Percent Identity: 73.050

alignment_block:

US-09-742-690-1 x AAY08989 ..

Align seg 1/1 to: AAY08989 from: 1 to: 409

12 CAGGTGACGTCAGGAGTCAGGAGGAGTGGTCCAGGCTGGGGGTC 61
 |||||
 263 GlnValGlnLeuGlnSerGlyGlyLeuValGlnProGlyGlySe 279
 |||||
 62 TGTGAGACTCTCTGTGAGCCTCGGGAGCGGCCACCGAGTGTGTC 111
 |||||
 279 rleuargLeuSerCysAlaAlaSerGlyLeuThrLeuThr.....T 293
 |||||
 112 ACTATGTTATGGCTGGTTCGCCAGGTTCCAGGAAGAGCGTGAGTT 161
 |||||
 293 hrTySerThrGlyTrpPheArgGlnAlaProGlyLysGluArgGluPhe 309
 |||||
 162 GTCGACGCTATTAGTGGAGT...GGTAAGAGACATGGTATAGACTC 208
 |||||
 310 ValGlyMetLeuGlyTrpSerGlyGlyGlyAsnThrTyTrAlaAspSe 326
 |||||
 209 CGTGAAGGCCGATTCACCATCTCCAGAGATAACGCCAAGACTACGGTT 258
 |||||
 326 rVallyGlyArgPheThrIleSerArgAspAsnAlaLysAsnMetValP 343
 |||||
 259 ATCTCAATGAACAGCCCTGAACCTGAGATACGGCCGCTTTATTATTG 308
 |||||
 343 heLeuGlnMetSerSerLeuLysProGluAspThrAlaValTyTrCys 359
 |||||
 309 GCCGTCGACCGCTGGCTGGATGATATTCCCTGCCGTTGGTTTGA 358
 |||||
 360 AlaAlaArgGlnProTyArgGlySerTyTrSerAspProAsnAsnTyRhi 376
 |||||
 359 CTACTGGGGCCAGGGACCCAGGTCCACGCTCCTCAGGATCTCATCAC 408

|||||
 376 sTyTrpGlyGlnGlyThrGlnValThrValSerSerGlySerHisHisH 393
 |||||
 409 ATCACCATCAGGATCCACCTCC 431
 |||||
 393 IsHisHisHisGlySerGlySer 400

seq_name: /SID55/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:AA199987

seq_documentation_block:

ID AAY08987 standard; Protein; 146 AA.

XX
 XX AC
 XX AAY08987;
 XX
 DT 20-AUG-1999 (first entry)
 XX
 DE Llama protein construct pPIC.HCV24-His2t.

XX Analytical device; electrically conductive solid support; immobilisation;
 KW electroactive; analysis; peptide hormone; steroid; disease marker; HCV24;
 KW diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama.
 XX Lama guanicoe glama.
 OS Synthetic.
 OS
 XX WO9927356-A1.
 XX
 PD 03-JUN-1999.
 XX
 XX 23-NOV-1998; 98WO-GB03495.
 PF
 XX
 XX 21-NOV-1997; 97EP-0309425.
 PR
 XX (UNIL) UNILEVER NV.
 PA (UNIL) UNILEVER PLC.
 PA
 XX Badley RA, Porter RA;
 XX WPI; 1999-385228/32.
 DR
 XX Analytical device including electrochemically active compound on
 PT solid support
 PT
 XX Example 6.2; Page 68; 78pp; English.

XX This invention describes a novel component for an analytical device
 CC comprising an electrically conductive solid support on which a compound
 CC (I) is immobilized which has an electroactive part (EA) with an
 CC electrochemical property that can be modulated in a detectable manner by
 CC binding it to a specific binding partner (II). Compounds used in the
 CC method of the invention include those with the formula R = hydrogen,
 CC hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally
 CC substituted), halo, amido, or amino; optionally one or more positions on
 CC the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or
 CC alkoxy (all optionally substituted), acid groups (organic or inorganic),
 CC halo, amido or amino. Devices that contain the component of the invention
 CC are used for qualitative and quantitative analysis of e.g. nucleic acid,
 CC hormones (peptide or steroid), disease markers, diagnostic indicators
 CC etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices
 CC can be used with turbid solutions (whole blood, milk, culture liquids
 CC etc.), unlike conventional colourimetric methods. This sequence
 CC represents a protein construct pPIC.HCV24-His2t which contains the llama
 CC HC-v fragment, HCV24 which is used in the method of the invention.
 XX

SQ Sequence 146 AA;

alignment_scores:
 Quality: 508.50 Length: 140
 Ratio: 4.422 Gaps: 1
 Percent Similarity: 82.143 Percent Identity: 72.857
 alignment_block:

```
US-09-742-690-1 x AAY08987
Align seg 1/1 to: AAY08987 from: 1 to: 146

12 CAGGTGACGTCAGGAGTCAGGGGAGGATTGGTCAGGCTGGGGCTC 61
1 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyGly 17
62 TCTGAGACTCTCTGTGACGCTCGGAGCGCCACCATGTCATGGTC 111
17 rLeuArgLeuSerCysAlaAlaSerGlyArgThrPheSer.....V 31
112 ACTATGTTGGCTGCTCCGCGAGTTCACGGGAGGCGGTGAGTTT 161
31 alTyAlaValGlyThrPheArgGlyGlyAlaProGlyLysGluArgGluPhe 47
162 GTGCGAGCTATTAGTGGAGTGTAAAGACATGATGTTAAAGACATCCGT 211
48 ValGlyTyPheGlyThrArgGlyGlyArgThrTyTyrAlaAspSerVa 64
212 GAAGGCCGATTCACCATCTCCAGATATACGGCAAGACTACGGTTATC 261
64 lLysGlyArgPheThrIleAlaIleAspAsnAlaLysAsnThrValTyrL 81
262 TGCAAATGAACACCTGAACCTGAAGATACGGCCGCTTTATTATTGTC 311
81 euGlnMetAsnSerLeuLysLeuAspAspThrAlaValTyrCysAla 97
312 GCTCGACCGTCCGCGTGGATGATATTTCCCTCCGCGTTCGGTTTGA 361
98 ValArgMetProTyrSerGlyAspTyrArgSerSerGlyThrTyrAsp 114
362 CTGGGCCAGGGACCGGTCACCGTCCTCCAGGATCTCATCACCATC 411
114 rTrpGlyGlnGlyThrGlnValThrValSerSerGlySerHisHisH 131
412 ACCATCAGGATCCACTCC 431
131 ishHisGlySerGlySer 137

seq_name: /SID55/gcgdata/geneseq/geneseq-emb1/AAL999.DAT: AAY08990

seq_documentation_block:
ID AAY08990 standard; Protein; 408 AA.
AC AAY08990;
XX
DE 20-AUG-1999 (first entry)
XX
DE Llama protein construct pPIC.scfV4155-link-HCV24-His2t.
XX
KW Analytical device; electrically conductive solid support; immobilisation;
KW electroactive; analysis; peptide hormone; steroid; disease marker; HCV24;
KW diagnostic indicator; antibody; bacteria; virus; fungi; yeast; llama.
XX
OS Lama guanicoe glama.
OS Synthetic.
XX
PN WO9927356-A1.
XX
PD 03-JUN-1999.
XX
PF 23-NOV-1998; 98WO-GB03495.
XX
PR 21-NOV-1997; 97EP-0309425.
XX
PA (UNIL ) UNILEVER NV.
PA (UNIL ) UNILEVER PLC.
XX
PI Badley RA, Porter RA;
XX
DR WPI; 1999-385228/32.
XX
```

PT Analytical device including electrochemically active compound on
XX solid support
PS Example 6.4; Page 72-73; 78pp; English.
XX This invention describes a novel component for an analytical device
CC comprising an electrically conductive solid support on which a compound
CC (1) is immobilized which has an electroactive part (EA) with an
CC electrochemical property that can be modulated in a detectable manner by
CC binding it to a specific binding partner (II). Compounds used in the
CC method of the invention include those with the formula R - hydrogen,
CC hydroxy, 1-14C alkyl, aryl, alkenyl or alkoxy (all optionally
CC substituted), halo, amido, or amino; optionally one or more positions on
CC the heteroaromatic rings are substituted by alkyl, aryl, alkenyl or
CC alkoxy (all optionally substituted), acid groups (organic or inorganic),
CC halo, amido or amino. Devices that contain the component of the invention
CC are used for qualitative and quantitative analysis of e.g. nucleic acid,
CC hormones (peptide or steroid), disease markers, diagnostic indicators
CC etc., e.g. antibodies, bacteria, viruses, fungi or yeast. These devices
CC can be used with turbid solutions (whole blood, milk, culture liquids
CC etc.), unlike conventional colourimetric methods. This sequence
CC represents a protein construct pPIC.scfV4155-link-HCV24-His2t which
CC contains the llama HC-V fragment, HCV3 which is used in the method of
CC the invention.
XX Sequence 408 AA;
SQ

alignment_scores:
Quality: 508.50 Length: 140
Ratio: 4.422 Gaps: 1
Percent Similarity: 82.143 Percent Identity: 72.857

alignment_block:
US-09-742-690-1 x AAY08990 ..
Align seg 1/1 to: AAY08990 from: 1 to: 408

12 CAGGTGACGTCAGGAGTCAGGGGAGGATTGGTCAGGCTGGGGCTC 61
263 GlnValGlnLeuGlnGluSerGlyGlyGlyLeuValGlnAlaGlyGly 279
62 TCTGAGACTCTCTGTGACGCTCGGAGCGCCACCATGTCATGGTC 111
279 rLeuArgLeuSerCysAlaAlaSerGlyArgThrPheSer.....V 293
112 ACTATGTTGGCTGCTCCGCGAGTTCACGGGAGGCGGTGAGTTT 161
293 alTyAlaValGlyThrPheArgGlyGlyAlaProGlyLysGluArgGluPhe 309
162 GTGCGAGCTATTAGTGGAGTGTAAAGACATGATGTTAAAGACATCCGT 211
310 ValGlyTyPheGlyThrArgGlyGlyArgThrTyrTyrAlaAspSerVa 326
212 GAAGGCCGATTCACCATCTCCAGATATACGGCAAGACTACGGTTATC 261
326 lLysGlyArgPheThrIleAlaIleAspAsnAlaLysAsnThrValTyrL 343
262 TGCAAATGAACACCTGAACCTGAAGATACGGCCGCTTTATTATTGTC 311
343 euGlnMetAsnSerLeuLysLeuAspAspThrAlaValTyrCysAla 359
312 GCTCGACCGTCCGCGTGGATGATATTTCCCTCCGCGTTCGGTTTGA 361
360 ValArgMetProTyrSerGlyAspTyrArgSerSerGlyThrTyrAsp 376
362 CTGGGCCAGGGACCGGTCACCGTCCTCCAGGATCTCATCACCATC 411
376 rTrpGlyGlnGlyThrGlnValThrValSerSerGlySerHisHisH 393
412 ACCATCAGGATCCACTCC 431
393 ishHisGlySerGlySer 399